普通高等学校教材

地理信息系统

刘 南 刘仁义 编著



高等教育出版社





ISBN 7-04-011237-X
9 787040 112375 >
定价 28.10 元

普通高等学校教材

地理信息系统

刘　南　刘仁义　编著

高等教育出版社

内容提要

本书系作者基于多年教学科研之积累，特别是结合作者领导的研发团队近年来对新一代GIS技术研发攻关的体验，同时参阅当今国内外众多同类教材及专著，编写而成。作者力求着眼于当前地理信息科学新视野来诠释基本概念，力求做到基础内容与高新内容之间的统一。全书分七章，主要包括：地理信息系统概述、空间数据组织与计算机表达、地理空间数据管理、地理空间数据处理、空间分析、地图制图及输出、GIS新技术发展。

该书着力在深入浅出上下功夫，教材中每一个新出现的基本概念都有足够的解释，以真正做到“循序渐进”。

该书不仅是一本适用于地学、测绘、海洋、农林等相关学科的不可多得的本科教材，研究生们、GIS或有关的专业人员阅读后，也会感到颇有收益。

图书在版编目（CIP）数据

地理信息系统/刘南，刘仁义编著. —北京：高等教育出版社，2002. 8

ISBN 7-04-011237-X

Ⅰ. 地… Ⅱ. ①刘… ②刘… Ⅲ. 地理信息系统-高等学校-教材 Ⅳ. P208

中国版本图书馆CIP数据核字（2002）第062740号

地理信息系统

刘南　刘仁义　编著

出版发行	高等教育出版社	购书热线	010-64054588
社　　址	北京市东城区沙滩后街55号	免费咨询	800-810-0598
邮政编码	100009	网　　址	http://www.hep.edu.cn
传　　真	010-64014048		http://www.hep.com.cn
经　　销	新华书店北京发行所		
排　　版	高等教育出版社照排中心		
印　　刷	北京未来科学技术研究所有限责任公司印刷厂		
开　　本	787×960　1/16	版　　次	2002年10月第1版
印　　张	22.5	印　　次	2002年10月第1次印刷
字　　数	420 000	定　　价	28.10元

前　言

当代计算机和信息科学技术的更新速度明显高于其他学科，而且发展得愈来愈快。回顾近十几年的计算机发展，人们难以想象十几年后，计算机、信息科学技术及相关领域会发展到什么程度。地理信息系统（geographic information systems，简称 GIS），或地理信息科学（geographic information science，可简写为 GIS、GISci 或 GIScience），是地理、测绘等地学与计算机、信息科学相结合的产物，它既是地学或地理学的一个现代化分支，也可以说是计算机和信息科学具有地理空间涵盖的分支。由于地理空间数据（geospatial data）或 GIS 的特殊复杂性，GIS 在为计算机和信息科学作贡献的同时，总是如饥似渴地、迅速地吸取计算机主流技术的各种最新成果，成为计算机科技领域中应用研究技术内容最丰富的分支之一。例如，当前的 GIS 技术，除自身特色以外，还包含图形图像、数据库、互联网及网络、组件、办公自动化、虚拟现实、人工智能等计算机技术新内容。计算机和信息科学发展有多快，GIS 的发展就有多快。20 世纪 80 年代甚至 90 年代前期的 GIS，现在已常常被称为“传统”的 GIS。

高速度的发展和更新，给计算机和 GIS 这些高科技领域的教学科研带来了新的特点。这里的专业人员对知识更新的投入特别大，写这一领域的教材，也与写传统领域教材有明显不同。十几年前，作者写《地球概论》（也是高教社出版的教材）等书时，很多内容可说是精雕细刻。现在不同了，高科技领域教材的某些内容可能很快过时；高新内容又太多，在不增加篇幅的条件下难与基础内容相兼顾；这里的专业人员又通常很繁忙，项目多，能专注写书的时间有限，致使在一些高科技书籍中，常常可见“匆忙草就”的痕迹。

作者也面临这样的问题。但是，作者不愿，也不敢有所懈怠，既然接受了编写任务，就必须很好完成，无论从内容上，还是形式上，所应当尽力达到任何一本好教材所应当达到的基本要求。因此，针对本教材读者对象的特点，经反复实践和思索，除注意结构、层次和文字等一般要求外，作者还力求做到：

在内容方面：第一，力求基础与高新内容之间的统一。GIS 领域近年来新发展内容很多，本书必须尽可能反映；但是，本书是基础课程教材，必须注重基本概念。为了在不增加篇幅的条件下兼顾二者，最好的做法只能是：着眼于当前地理信息科学新视野来诠释基本概念，并着力在深入浅出上下功夫。为此，本

书中有不少地方采用新鲜的写法或独特的提法,力求将一些常易混淆的概念和GIS新发展所涉及的内容,变得较为清晰和较容易理解。当然,不少地方基于作者自己的理解,文责自负。

第二,对估计不会很快过时的基础概念,多花笔墨和功夫,写出循循善诱的教材特点;当然,也不忽略其他内容。整个第七章和其他各章的部分(用小字印出)有关的内容,不作为正文和基本要求。

第三,不提前使用尚未讲解的、学生难懂的概念(这是写书最难做到的一点),本书正文中每一个新出现的基本概念都有足够的解释,以真正做到"循序渐进"。

第四,在适当的地方交代教材的写作思路,亦即教和学的思路。

形式为内容服务。在形式方面,本书也采取了一些便利于阅读的措施。例如:

第一,避免过多的标题层次。标题层次过多,例如章、节之下的一、(一)、1、(1)、①、A等,往往使读者看不清内容的层次和逻辑关系。十余年前,华东师大已故金祖孟教授认为,一般教材只保留章、节、分节和小节四个标题层次为宜,嵌套关系很深的内容可以通过疏散章节的途径来解决。作者赞同这种观点。本书中,节和分节采用多重数字序号,如§2.2为第二章二节,2.2.3为第二章二节第3分节。但第四级(小节)的标题序号采用单纯数字,如1、2、3等,且标题名称给出该小节的段落大意。

第二,章节的长短尽可能均匀,大部分的节和分节所含页数差别不大。

第三,章节所包含的子部分数目也尽可能匀调,通常一章包含3~4节,一节包含3~4分节,一分节包含3~4小节等。

第四,控制自然段落包含的行数,避免使读者感到劳累的大块、超大块自然段落。

第五,尽可能做到语句读起来比较"上口"。

作者深感庆幸的是,最近几年国内有很多专门论述GIS的好书推出,作者从中受益匪浅。其中特别值得感谢的是:陈述彭、鲁学军、周成虎的《地理信息系统导论》,龚健雅的《地理信息系统基础》,邬伦、刘瑜、张军等的《地理信息系统——原理、方法和应用》等。早期的GIS书籍,如张超、陈丙咸、邬伦的《地理信息系统》、黄杏元的《地理信息系统概论》和彭望琭的《遥感数据的计算机处理与地理信息系统》等,作者也在此一并谢之。本书实际上也是作者所在的整个浙江省GIS重点实验室团队共同劳动的结晶,在这里,我们要特别感谢严志明、陆丽珍、谢炯、高锡章等博士和硕士研究生。本书的编写得到高等教育出版社的黎勇奇先生、徐丽萍女士的热情指导和帮助,特别是靳剑辉先生为此付出大量的精力,提出非常宝贵的指导意见,在此表示衷心的感谢。

本书编写的具体分工是:刘南编写第1、3章的第1和第2节,以及第2、4和5章的第1节,其余章节由刘仁义编写(约总字数的75%);全书最后由刘南统稿。

因时间紧,疏漏难免,习题集等教学材料也来不及配套。种种不足,恭请指正和原谅。

编著者　2002年5月于浙江大学

第一章 地理信息系统概述 …………………………………………… (1)
§1.1 数据、信息和信息系统 …………………………………………… (1)
1 1 1 系统 …………………………………………… (1)
1 1.2 数据和信息 …………………………………………… (2)
1 数据 …………………………………………… (2)
2. 信息 …………………………………………… (2)
3. 数据和信息的关系 …………………………………………… (2)
1 1.3 信息系统 …………………………………………… (2)
1. 信息系统的本质特征 …………………………………………… (2)
2. 基于计算机的信息系统 …………………………………………… (3)
3. 信息系统的构成要素 …………………………………………… (3)
4. 信息和信息系统的层次性 …………………………………………… (4)
5. 数据挖掘和知识发现 …………………………………………… (5)
§1.2 地理信息系统基本概念和本质特征 …………………………………………… (5)
1.2.1 地理信息系统定义 …………………………………………… (6)
1 2 2 地理空间数据和信息 …………………………………………… (6)
1 地理空间数据和信息 …………………………………………… (6)
2. 地理空间数据和信息的三个基本特征 …………………………………………… (7)
3. 地理空间数据区别于一般计算机数据的本质特征 …………………………………………… (8)
4. 有关地理空间数据和信息的用词问题 …………………………………………… (9)
1.2.3 地理信息系统区别于一般信息系统的主要特点 …………………………………………… (9)
1 地理空间数据和信息的特殊复杂性 …………………………………………… (10)
2. 必须具备科学可视化功能 …………………………………………… (10)
3 区域性和多层次 …………………………………………… (11)
4 数据量较大 …………………………………………… (13)
5 注重空间分析 …………………………………………… (13)
1 2 4 地理信息系统与计算机科学技术的关系 …………………………………………… (13)
1. 地理信息系统与计算机科学技术总体的关系 …………………………………………… (13)

2 地理信息系统与各计算机具体分支技术的关系 …… (14)
3 地理信息系统与一般计算机制图系统的关系和区别 …… (15)
§1.3 地理信息系统应用和发展 …… (16)
1.3.1 地理信息系统的社会需求、产生与发展 …… (17)
1 导致地理信息系统产生的社会条件 …… (17)
2 地理信息系统产生和发展 …… (18)
3 地理信息系统在我国的发展 …… (19)
1.3.2 地理信息系统的应用领域 …… (20)
1 85%～90%的政府部门 …… (20)
2 企业和商用领域 …… (22)
3 民用 …… (22)
1.3.3 地理信息系统与地理信息科学 …… (23)
1 地理信息科学的形成 …… (23)
2 地理信息科学的涵盖 …… (23)
3 有关新名词讨论 …… (24)
4 本书的定位 …… (25)
§1.4 地理信息系统的功能及组成 …… (25)
1.4.1 地理信息系统的主要任务 …… (25)
1 某个地方有什么 …… (25)
2. 符合某些条件的实体在哪里 …… (26)
3. 变化趋势 …… (26)
4 空间结构特征 …… (26)
5. 模拟 …… (27)
1.4.2 地理信息系统的主要功能 …… (27)
1. 地理信息系统的基本功能 …… (27)
2. 空间分析与模型分析功能 …… (29)
1.4.3 地理信息系统的主要组成成分 …… (30)
1. 地理空间数据和信息 …… (30)
2 硬件系统 …… (30)
3 软件系统 …… (31)
4 系统开发、管理和使用人员 …… (31)
1.4.4 地理信息系统的软硬件组成 …… (32)
1 GIS 硬件系统的特点 …… (32)
2 输入、输出设备 …… (33)
3 GIS 软件系统 …… (33)

4 地理信息系统中的"上层建筑" …… (34)

第二章 空间数据组织与计算机表达 …… (37)
§2 1 从地理空间现象到计算机世界 …… (37)
2 1 1 从现实世界到计算机世界 …… (37)
1 模型与概念模型 …… (37)
2 数据模型 …… (38)
3 数据结构 …… (38)
4 文件格式 …… (39)
2 1 2 GIS 中地理空间数据组织的主要对象 …… (39)
1 数值计算领域的地理空间数据组织 …… (39)
2 再谈地理信息科学之涵盖 …… (40)
3 GIS 所抽象、表达的地理事物和现象 …… (41)
2.1 3 空间对象和空间关系 …… (42)
1. 空间对象 …… (42)
2 空间关系 …… (43)
3. 拓扑空间关系 …… (44)
2.1 4 有关地理空间数据结构和模型学习的几点总体说明 …… (45)
1. 关于空间数据结构和空间数据模型两个概念之间的关系 …… (45)
2. 关于属性数据的组织 …… (45)
3. GIS 可视化要求和传统美术的两种基本技法 …… (47)
4 两种基本空间数据模型或结构，及其与计算机图形图像的关系 …… (48)
§2.2 矢量空间数据模型及结构 …… (48)
2 2.1 矢量空间数据模型 …… (48)
1. 二维空间坐标系 …… (49)
2 几何数据 …… (49)
3 属性数据 …… (50)
4 唯一标识符 …… (50)
2 2.2 有代表性的矢量空间数据结构 …… (51)
1 Spaghetti 结构 …… (51)
2. 拓扑空间数据结构 …… (52)
3. 拓扑关系及其表达 …… (53)
4 讨论 …… (54)
2 2.3 不同格式的拓扑结构数据实例 …… (55)
1 不同格式的拓扑结构数据 …… (55)

2 DLG 数据格式及其实例 …… 55
3 Arc Info 的拓扑数据表达 …… (57)
4 再谈数据结构和文件(或数据)格式 …… (58)
2.2.4 矢量结构的数据的输入、编辑和输出 …… (58)
1 矢量结构的数据的获取和输入 …… (58)
2 矢量结构数据的前处理和编辑 …… (60)
3 一般矢量数据的前处理和编辑 …… (60)
4 拓扑结构数据的前处理和编辑 …… (61)
5 矢量结构的数据的输出 …… (62)
§2.3 栅格空间数据模型及结构 …… (63)
2.3.1 栅格空间数据模型 …… (63)
1 规整格网 …… (63)
2. 二维空间坐标系 …… (64)
3 属性数据和分辨率 …… (65)
4. 矩阵数组 …… (66)
2.3.2 完全栅格空间数据结构 …… (66)
1. 数据量及其与精度的关系 …… (67)
2 不同地学应用类型的栅格数据 …… (67)
3 多重属性下的栅格数据结构 …… (69)
4 多层栅格数据结构的文件格式 …… (70)
2.3.3 压缩的栅格空间数据结构 …… (70)
1 游程长度编码或块码 …… (70)
2 链码 …… (71)
3. 四叉树编码 …… (71)
4. 地学栅格图像压缩不同于一般图像压缩的特点 …… (73)
2.3.4 栅格数据的采集、输入和输出 …… (74)
1 栅格数据结构在 GIS 输入和输出上的优势 …… (74)
2 栅格数据的采集、输入 …… (75)
3 栅格数据的输出 …… (76)
§2.4 基于两种数据结构的进一步讨论 …… (77)
2.4.1 矢量和栅格数据结构的比较分析 …… (77)
1 矢量和栅格数据结构总体比较 …… (77)
2 矢量和栅格数据结构具体优缺点比较 …… (77)
3 矢量和栅格数据结构的发展 …… (78)
4 矢量和栅格数据结构的选择 …… (79)

2.4.2 栅格、矢量数据结构的相互转换 ………………………… (80)
1. 矢量数据向栅格数据转换的主要步骤 ………………………… (80)
2. 点状、线状和面状空间对象矢量数据的栅格化 ………………… (80)
3. 栅格格式的遥感分类图像向矢量数据的转换 ………………… (81)
4. 扫描地图的栅格图像向矢量数据的转换 ……………………… (82)
2.4.3 其他空间数据结构 ……………………………………………… (83)
1. 镶嵌数据结构 …………………………………………………… (83)
2. 不规则三角网数据结构 ………………………………………… (85)
3. 栅格矢量混合结构 ……………………………………………… (86)
4. 三维空间数据结构 ……………………………………………… (87)
5. 简短的小结 ……………………………………………………… (88)

第三章 地理空间数据管理 ……………………………………………… (89)
§3.1 数据管理基础知识 ……………………………………………… (89)
3.1.1 数据库及其管理系统的基本概念 ………………………………… (90)
1. 数据组织和管理的层次 ………………………………………… (90)
2. 数据库 …………………………………………………………… (90)
3. 数据库管理系统 ………………………………………………… (91)
3.1.2 数据管理技术的发展：文件系统与数据库系统 ………………… (92)
1. 数据管理技术的发展 …………………………………………… (92)
2. 数据库系统的进一步发展 ……………………………………… (93)
3. 文件管理系统 …………………………………………………… (94)
3.1.3 关系数据库系统技术 ……………………………………………… (95)
1. 传统数据库系统技术的发展 …………………………………… (95)
2. 关系数据库技术 ………………………………………………… (96)
3. 关系模型的局限和不足 ………………………………………… (98)
3.1.4 面向对象数据库系统 ……………………………………………… (98)
1. 新一代数据库技术 ……………………………………………… (98)
2. 面向对象数据模型 ……………………………………………… (99)
3. 面向对象数据库系统 …………………………………………… (101)
4. 面向对象数据库系统的实现方式 ……………………………… (101)
§3.2 地理空间数据管理 ……………………………………………… (102)
3.2.1 地理空间数据管理不同于常规数据管理的特点 ………………… (102)
1. 栅格地理空间数据的管理 ……………………………………… (102)
2. 矢量地理空间数据的管理 ……………………………………… (103)

3 空间数据管理具有明显的海量、区域性和多层次特点 …………（104）
4 空间数据管理涉及到较复杂的实体类型和空间操作 …………（104）
3.2.2 传统的地理空间数据库系统 …………（105）
1 地理空间数据库系统 …………（105）
2 传统的地理空间数据库系统技术及其发展 …………（106）
3 传统GIS数据库系统技术，关系数据库与文件系统的混合管理 …………（107）
3.2.3 地理空间数据分区域、分层次的组织管理 …………（108）
1 地理空间数据分幅（分区）分层的组织管理 …………（108）
2 工作层和工作区 …………（109）
3 主要GIS软件的工作层和工作区 …………（110）
4 图库管理和空间索引 …………（111）
5 属性文件的组织 …………（112）
3.2.4 近年来GIS数据库技术的发展及展望 …………（113）
1. 空间数据库引擎 …………（113）
2. 目前空间数据库引擎的两种主要方式 …………（114）
3 面向对象空间数据库管理系统 …………（115）
4 其他发展趋势 …………（115）
§3.3 地理空间数据的元数据 …………（115）
3.3.1 元数据一般概念 …………（115）
1 元数据的由来 …………（115）
2 元数据概念 …………（116）
3. 元数据的分类 …………（117）
4 元数据的管理 …………（117）
3.3.2 地理空间数据元数据的概念和标准 …………（118）
1 地理空间数据元数据及意义 …………（118）
2 地理空间数据元数据的常用具体概念 …………（119）
3 空间数据元数据的标准 …………（121）
3.3.3 几种地理空间数据元数据标准及实例 …………（121）
1 美国FGDC元数据标准 …………（121）
2 ISO/TC 211的元数据标准草案 …………（123）
3 中国基础地理信息元数据实例 …………（124）
3.3.4 空间数据元数据的获取和应用 …………（125）
1 空间数据元数据的获取过程 …………（126）
2 空间数据元数据的获取方法 …………（126）
3. 空间数据元数据的应用及其意义 …………（126）

第四章 地理空间数据处理 ……………………………… (128)
§4.1 有关地理空间数据处理的总体说明 ……………………………… (128)
4.1.1 地理空间数据处理总体说明的必要性及内容 ……………………………… (128)
1. 地理空间数据处理总体说明的必要性 ……………………………… (128)
2. 本节的内容和做法 ……………………………… (129)
3. 广义的地理空间数据处理的概念框架 ……………………………… (130)
4.1.2 非GIS特色的地理空间数据处理 ……………………………… (132)
1. 普通数据处理 ……………………………… (132)
2. GIS中的一般的图形处理 ……………………………… (132)
3. GIS中的一般的图像处理 ……………………………… (134)
4.1.3 GIS特色的数据处理和广义的空间数据处理 ……………………………… (135)
1. GIS特色的数据编辑 ……………………………… (135)
2. 空间数据质量与精度控制 ……………………………… (136)
3. 其他GIS特色的数据处理 ……………………………… (138)
4. 广义的空间数据处理 ……………………………… (138)
§4.2 地理坐标变换及有关问题 ……………………………… (139)
4.2.1 地球表面形状和地理坐标 ……………………………… (139)
1. 地球表面形状的几级近似表达 ……………………………… (139)
2. 两种地球椭球体 ……………………………… (141)
3. 三种地理坐标 ……………………………… (142)
4.2.2 地图投影和坐标转换 ……………………………… (143)
1. 地图投影 ……………………………… (143)
2. 高斯-克吕格坐标 ……………………………… (144)
3. 平面地图直角坐标间的转换:坐标校正 ……………………………… (148)
4.2.3 其他有关问题 ……………………………… (150)
1. 关于坐标校正的讨论 ……………………………… (150)
2. 引进国外GIS软硬件带来的有关问题 ……………………………… (151)
3. 计算机内的比例尺与坐标单位 ……………………………… (152)
§4.3 地理空间数据插值 ……………………………… (154)
4.3.1 基本概念 ……………………………… (154)
1. 空间数据插值的概念和必要性 ……………………………… (154)
2. 空间变量分布与不同空间分布的数据源 ……………………………… (155)
3. 空间数据插值基本思想和理论假设 ……………………………… (156)
4. 整体插值法和局部插值法 ……………………………… (157)

4.3.2 整体插值法 …… (157)
1. 变换函数插值 …… (157)
2. 趋势面分析和其他数学模拟 …… (158)
3. 边界插值方法 …… (159)
4.3.3 采用活动窗口的局部插值法 …… (160)
1. 基本方法和技术路线 …… (160)
2. 局部函数法 …… (161)
3. 移动平均法 …… (162)
4. 搜寻策略 …… (163)
4.3.4 其他局部插值法 …… (163)
1. 克里金插值 …… (163)
2. 加密插值和配准插值 …… (165)
3. 样条函数 …… (166)
§4.4 地理空间数据的三维处理 …… (167)
4.4.1 空间数据三维处理的基本概念 …… (167)
1. 空间数据三维处理的两种基本思路 …… (167)
2. 三维地形表达的传统方式 …… (168)
3. GIS三维处理的主要内容,DEM和DTM概念 …… (169)
4. 现代三维地形表达的5个层次 …… (171)
4.4.2 数字高程模型及其生成 …… (173)
1. 数字高程模型生成的一般思路 …… (173)
2. 高程矩阵的特点及其生成 …… (174)
3. TIN模型及其生成 …… (174)
4. 关于TIN生成方法的进一步讨论 …… (176)
4.4.3 三维可视化表达的某些技术步骤 …… (177)
1. 数字高程模型的可视化 …… (177)
2. 一定观察角度鸟瞰的透视立体图 …… (178)
3. 一定观察角度、一定光源条件下的灰度或彩色立体图 …… (179)
4. 进一步的三维表达技术 …… (181)
4.4.4 地形派生数据或地形分析 …… (183)
1. 坡度和坡向 …… (183)
2. 等值线 …… (183)
3. 视线图 …… (184)
4. 地形轮廓线及其他 …… (184)

第五章 空间分析 (186)

§5.1 空间分析一般概念与空间查询 186

5.1.1 空间分析概念 186

1. 关于"空间分析"概念的总体说明 (186)
2. 本书的做法及空间分析概念框架 (187)
3. 关于上述框架的说明 187

5.1.2 基本空间查询 189

1. 关于"基本空间查询"的概念 (189)
2. 从地物空间位置(或图形)特征查询属性 (190)
3. 从属性查询空间位置(或图形)特征 (190)
4. 用户给定的几何窗口查询 (191)

5.1.3 较复杂的空间查询 (192)

1. 空间关系查询 (192)
2. 与统计分析交互的空间查询 (194)
3. 虚拟现实环境下的空间查询 194

§5.2 程式化空间分析 (195)

5.2.1 空间操作及分析 (196)

1. 空间操作的概念 196
2. 空间剪裁与空间筛选 (197)
3. 多边形合并 (198)
4. 在面向对象的数据模型基础上对空间对象的空间操作 (199)

5.2.2 叠置操作及分析 (199)

1. 点与多边形叠置 (199)
2. 线与多边形叠置 (200)
3. 多边形叠置 (200)
4. 栅格 GIS 的叠置分析 (202)

5.2.3 缓冲区分析 (203)

1. 缓冲区分析的概念 (203)
2. 缓冲区的生成 (204)
3. 缓冲区分析实例 (205)

5.2.4 网络分析 (205)

1. 网络与网络分析的概念 205
2. 最佳路径 (206)
3. 资源分配 (208)
4. 服务中心选址 (209)

§5.3 其他空间分析 …… (210)
5.3.1 空间统计分析及其与栅格数据处理方法的联系 …… (210)
1. 空间统计分析的内容和意义 …… (210)
2. 空间统计分析与栅格数据处理方法的联系 …… (212)
3. 空间变量筛选 …… (212)
4. 地物分类 …… (214)
5. 综合分析评价 …… (215)
5.3.2 分析地图学与空间分析 …… (217)
1. 基于矢量地图的空间统计分析 …… (217)
2. 基于地理空间和距离拓展概念所进行的空间分析 …… (218)
3. 分析地图学简介 …… (219)
5.3.3 进一层的空间分析 …… (220)
1. 多种基本空间操作或分析功能的组合 …… (220)
2. 应用领域的专业模型与GIS相结合 …… (220)
3. 空间区位-配置模型 …… (221)
4. 专家系统 …… (223)
5. 空间决策支持模型 …… (224)

第六章 地图制图及输出 …… (227)
§6.1 地图绘制与空间对象的符号化 …… (227)
6.1.1 GIS输出与地图绘制 …… (227)
1. 地理信息系统输出产品 …… (227)
2. 地图绘制、GIS输出与GIS可视化 …… (228)
3. 地图绘制:空间对象的符号化过程 …… (229)
4. GIS中的空间实体符号化过程的一般原理和步骤 …… (230)
6.1.2 地图符号及其分类 …… (231)
1. 点位符号 …… (231)
2. 线状符号 …… (232)
3. 面状符号 …… (233)
6.1.3 专题地图制图特点 …… (234)
1. 普通地图和专题地图 …… (234)
2. 专题地图分类 …… (235)
3. 专题地图内容表示方法 …… (236)
4. 专题地图的GIS实现 …… (238)
§6.2 地图符号制作或绘制 …… (239)

6.2.1 地图符号制作方法 …… (239)
1. 编程法 …… (239)
2. 直接信息法 …… (240)
3. 间接信息法 …… (241)
6.2.2 矢量地图点、线、面符号的具体绘制 …… (241)
1. 矢量点符号绘制 …… (241)
2. 矢量线符号绘制 …… (242)
3. 矢量面符号绘制 …… (243)
6.2.3 地图符号库及其应用 …… (244)
1. 地图符号库一般概念 …… (244)
2. 地图符号库的检索机制 …… (245)
3. 地图符号的二次开发 …… (246)
4. 现有软件的地图符号系统的应用特点 …… (246)
§6.3 地图注记、排版和输出 …… (248)
6.3.1 地图注记 …… (248)
1. 注记的数据结构 …… (248)
2. 注记方式与注记编辑 …… (249)
3. 地名自动注记 …… (250)
6.3.2 地图排版 …… (251)
1. 颜色配置 …… (251)
2. 图幅整饰 …… (253)
3. 地图排版布局 …… (254)
6.3.3 地图输出 …… (255)
1. 地图输出系统 …… (255)
2. 绘图仪或打印机输出 …… (255)
3. 自动制版输出 …… (255)
4. 电子地图制作 …… (256)

第七章 GIS 新技术发展 …… (259)
§7.1 万维网地理信息系统 …… (259)
7.1.1 Web GIS 概念 …… (260)
1. Web GIS 基本概念 …… (260)
2. Web GIS 组成 …… (260)
3. Web GIS 的应用及意义 …… (261)
7.1.2 Web GIS 主要特点 …… (261)

1 基于Internet/Intranet标准 …… (261)
2 分布式服务体系结构 …… (261)
3 高效利用空间数据资源 …… (262)
4 其他优点 …… (262)
7.1.3 Web GIS基本原理和构建方法 …… (263)
1 Web GIS基本原理 …… (263)
2 利用CGI技术方法构建Web GIS系统 …… (264)
3 利用服务器端应用程序接口建立Web GIS系统 …… (265)
4 利用plug-in插件技术方法建立Web GIS系统 …… (266)
5. 利用ActiveX控件和DCOM组件对象模型技术建立Web GIS系统 …… (267)
6. 利用Java编程语言建立Web GIS系统 …… (267)
7.1.4 主要Web GIS系统平台基本功能 …… (269)
1 ESRI公司的Internet Map Server(ArcIMS)平台 …… (269)
2. MapInfo公司的MapXtreme平台 …… (271)
3 Autodesk公司的Autodesk MapGuide平台 …… (274)
4 Intergraph公司的GeoMedia Web Map平台 …… (275)
5 武汉吉奥公司的GeoSurf平台 …… (277)
6. 国家遥感应用工程技术研究中心的地网GeoBeans平台 …… (278)
§7.2 虚拟现实地理信息系统 …… (279)
7.2.1 虚拟现实技术 …… (279)
1 虚拟现实技术基本概念 …… (279)
2. 虚拟现实技术的主要特征 …… (280)
3. 虚拟环境概念的四层含义 …… (280)
7.2.2 VR GIS基本概念及其特点 …… (281)
1 VR技术与GIS技术的结合 …… (281)
2. VR GIS基本概念及特点 …… (283)
3 VR GIS与三维地形可视化 …… (284)
7.2.3 三维显示工具OpenGL技术 …… (286)
1. OpenGL基本概念 …… (286)
2 OpenGL工作结构 …… (287)
3 OpenGL函数 …… (290)
4 OpenGL与三维地形虚拟现实 …… (290)
7.2.4 VR-GIS地学应用实例 …… (291)
§7.3 时态地理信息系统 …… (293)
7.3.1 时态GIS基本概念 …… (293)

1 地理空间数据和信息的时态特征 …… (293)
2. 时态地理信息系统 …… (293)
3 时态 GIS 与时态数据库 …… (294)
4 几种有代表性的时空数据模型 …… (295)
7.3.2 基态修正扩展模型 …… (296)
1 现时库 过程库和历史库 …… (296)
2 区段快速索引 …… (297)
3 变粒度存贮因子 …… (298)
7.3.3 基态修正扩展模型的应用实例 …… (299)
1 土地时空过程分析 …… (299)
2 土地时空数据结构 …… (300)
3. 区段快速索引及变粒度索引因子的实现 …… (302)
4 数据模型的拓扑关系及时空数据存贮 …… (302)
5. 土地产权产籍管理信息系统应用实例 …… (303)
§7.4 客户/服务器体系结构及空间数据库引擎 …… (305)
7.4.1 客户/服务器体系结构 …… (305)
1 客户/服务器基本概念 …… (306)
2. 典型两级客户/服务器开发模式及其局限性 …… (307)
3. 三级客户/服务器服务模式 …… (308)
7.4.2 空间数据库引擎 …… (309)
1 空间数据库引擎的概念和作用 …… (309)
2. ESRI 公司的 ArcSDE …… (311)
3 Oracle/MapInfo 公司的 Oracle Spatial …… (311)
7.4.3 自主开发空间数据库引擎 …… (313)
1. SDG 数据模型组织及操作 …… (313)
2 数据模型的空间图形要素表达 …… (314)
3. 数据模型主要技术条件 …… (315)
4 数据模型体系结构 …… (316)
5 SDG 应用实例 …… (317)
§7.5 组件式 GIS …… (317)
7.5.1 COM 基本概念及实现技术 …… (318)
1 COM 基本概念 …… (318)
2 OLE 技术概念 …… (319)
3 DCOM 基本概念 …… (320)
4 ActiveX 技术 …… (321)

7.5.2 组件式GIS基本概念及特点 …… (322)
1 组件式GIS的基本概念 …… (322)
2 传统GIS存在的问题 …… (322)
3 组件式GIS的特点 …… (324)
4 组件式GIS体系结构 …… (325)
7.5.3 几种主要组件式GIS平台及其比较分析 …… (327)
1 MapObjects平台 …… (327)
2 MapX平台 …… (328)
3 GeoMedia平台 …… (329)
4. 几种主要组件式GIS平台功能比较 …… (332)

参考文献 …… (335)

建议经常浏览的Web Site …… (339)

第一章
地理信息系统概述

§1.1 数据、信息和信息系统

地理信息系统是一种信息系统。本节介绍有关"信息系统"的若干基本概念，为进一步学习地理信息系统迈开第一步。

1.1.1 系统

系统科学和系统工程是现代科学技术的一项重要发展，现在很多领域都采用系统科学和系统工程综合集成的思想、方法和手段去分析、研究和解决问题。按照系统科学的观点，整个宇宙，包括天体系统、地球、社会、生物体等，都有层次、有目的、有组织地形成一个个系统。系统具有特定功能，保持动态的相对稳定，特定条件下发生突变；系统是对外开放的，在系统内部各成分之间，以及系统与其环境之间进行着物质、能量或信息的交换，也就是不断有输入和输出。系统通过各子成分之间，以及系统与其环境之间的相互协调，而获得最好的整体功能。图1-1是描述系统一般组成的简单示意图。

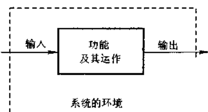


图1-1 系统一般组成示意

本书所关心的系统是信息系统或地理信息系统，它是一种特定类别的系统。信息系统是一种采集、输入数据或低级信息，按照人们的指令进行加工处理，提取、输出有用信息乃至知识的系统。可见，欲进一步理解信息系统，首先需要了解"数据"和"信息"的概念。

不同学科对信息系统的定义不尽相同。例如，从通讯或电信技术的角度，信息系统首先是传送信息的系统。本领域采用"处理数据、提取信息"的信息系统概念，与计算机科学和信息科学领域相同。

1.1.2 数据和信息

1. 数据

数据(data)和信息是信息系统的两个最基本概念。数据指人类在认识世界过程中,定性或定量描述认识目标的直接记录或原始资料。作为事件的直接记录,数据通常表现为可以识别的符号,包括数字、文字、符号、图形、图像等形式。数据具有多种多样的形式,可以由一种数据形式转换为其他数据形式。

2. 信息

信息(Information)是近代科学中一个具有重要应用意义的专门术语。狭义信息论认为,信息是人们获得信息前后对事物认识的差别;广义信息论认为,信息是指主体(人、生物或机器)与外部客体(环境或其他的人、生物或机器)之间的一切有用的消息或知识,是表征事物特征的一种普遍形式。不论从哪种角度,我们可以将信息理解为对人类认识有意义的消息或因素。

信息具有客观性、实用性、传输性和共享性。首先,信息是客观存在的,不随主观意志为转移;其次,信息有实用价值,是人类决策的重要依据;信息的传输性,是说信息可以在信息发送者和接受者之间传输、包括通过现代网络传输技术在不同信息系统、不同用户之间流转和交换;信息还具有共享性,它与实物不同,可以传输给多个用户,为多个用户共享,而其本身并无损失。信息的这些特点,使信息成为当代社会发展的一项重要资源。

3. 数据和信息的关系

信息与数据密不可分,因为数据是信息的载体,信息是潜在于数据中的意义。作为人类认识过程的原始素材,数据包含着对人类有意义的信息,且数据所包含信息的内容不因数据形式的转换而改变。但是,数据本身并不是信息,数据中蕴藏的信息通常不会自动地、明显地呈现出来;人们要从数据中得到信息,必须对数据作出解释,对数据进行处理(运算、排序、编码、分类、增强等),理解数据的含义,从而才能提取到数据中所包含的信息。

一言以蔽之,数据是人类认识过程的直接记录或原始素材,而信息是对数据的解释、是对数据加工后的有认识意义的结果。

1.1.3 信息系统

1. 信息系统的本质特征

信息系统正是采集、输入数据或低级信息,进行加工处理,输出有用信息的系统。信息系统的本质特征,就是从数据中提取和发掘有用信息。由此可见,一个信息系统的优劣应当根据它所提取信息的质量和数量来判断;而后者又常常取决于其分析功能和分析模型。一般而言,信息系统的分析功能愈强,分析

模型愈完善,能产生的信息质量就愈高,数量就愈多,信息系统就愈成功。

近年来,人们对信息系统提出了更高的、即"智能化"的要求,智能化的信息系统不仅应提取、输出信息,而且能提炼、发现知识。因此,图1.2中信息系统的输出端不像过去仅输出信息,而是"输出信息乃至知识"。这方面内容将在§1.4中进一步讨论。

从数据到信息乃至知识的过程,类似于从感性认识与理性认识的过程。信息系统从大量的数据中提炼出有用的信息乃至知识,做的就是一番去粗取精、去伪存真,由表及里,由此及彼的工作。

2. 基于计算机的信息系统

按照信息系统的定义或本质特征,从广义上说,很多高等动物的头脑,特别是人脑,都是很好的信息系统。不过,现在流行观念中的信息系统,通常指人机系统,即按照人们的意志或指令,进行数据加工处理(管理、分析、表达等),提取对人有用信息的计算机系统。众所周知,计算机的应用导致了一场信息革命,计算机科学与技术的突飞猛进,使应用计算机系统收集数据、并将数据处理成信息乃至知识成为可能。不仅如此,基于计算机的信息系统还具有不少优于人脑系统的数据处理功能,因而愈来愈成为人们认识和决策过程有力的辅助工具。

本书所说的信息系统,皆是指基于计算机的信息系统,对此下文将不再说明。

不过,并非所有的计算机系统都是信息系统。例如,很多单纯的文字处理或图形图像处理的计算机系统,不能算是信息系统,因为它们不能从数据处理中提取信息。

在各种计算机技术中,与信息系统关系较密切的技术首推数据库系统技术,因为数据库系统技术具有很好的管理、分析和处理数据的功能。一部分信息系统直接基于数据库系统,利用数据库系统的数据处理和分析功能。不仅如此,数据库还是所有信息系统的必备部分,信息系统要能按人们的需要生产信息,就必须具备存储有大量数据的仓库,并能对仓库内的数据操纵自如。如果将信息系统比做自动化加工厂,那么,数据库就是加工厂的管理有序的原料库,库中的原料有组织地加以分类和存放,能自动化地有效存取和利用。

从应用的角度,信息系统可分为多种类别,其中,管理信息系统(management information systems,常简称为MIS)和地理信息系统是常见的两种信息系统。不过,这两种信息系统概念有部分交叠,因为地理信息系统通常也具有一部分MIS特征。

3. 信息系统的构成要素

信息系统包括数据及信息、硬件、软件和人员四大要素。

数据及信息:数据是信息系统分析、处理的对象;获取信息是信息系统的目的。数据及信息构成信息系统的应用基础。

计算机硬件：是各类计算机处理及终端设备，包括各种输入 输出设备

计算机软件：是计算机程序系统，指挥各类硬件运作，支持数据和信息的采集、存储、加工和再现，以及回答用户问题等。

人员：包括开发人员或应用人员，他们是信息系统的建设者，或信息系统所服务的对象。

在一般有关书籍中，信息系统的四大要素通常写为数据、硬件、软件和用户。作者认为那种写法不够明确，因为信息系统以外的其他计算机系统一般也包含这四要素。为此，本书将“数据”要素改写为“数据及信息”，以表征信息系统的本质特点。此外，同其他计算机系统相比，信息系统更重视人的因素，因为信息系统是人们认识和决策过程的重要辅助工具，必须不断地改善其模拟和表达人脑认识和决策过程的能力；而且，人的知识、经验对数据处理有较大影响，信息系统获得信息的数量和质量，与开发或应用人员的知识水平密切相关。

4　信息和信息系统的层次性

信息有不同的层次。信息系统输出的信息，可能又作为数据被输入到这个或那个信息系统中，被进一步提炼，生成更进一层的信息；进一层产品还可能再送入信息系统提炼。例如，输入地理信息系统的数据，除野外观测数据、原始遥感数据、GPS 测量数据等第一手数据外，还常常是过去的信息产品，如数字化的地图、加工过的图像，以及整理出来的表格和统计数据等。地理信息系统将它们进一步处理加工，生成进一层的信息，如新的地图等。

可见，输入信息系统的，既可以是第一手数据，也可能是数据形式的、过去提取出的信息。为了区分这两种情况，有些书籍将这两种数据分别称为“原生数据”和“再生数据”，再生数据也称为“遗产”(legacy)数据。从信息的层次性角度，再生数据实际上是较低层次的信息(图 1－2)。

信息的层次性是一种客观存在。尽管有关书籍中对此一般没有明确的提法，作者认为，明确地讲述和分析信息的这一特征，有利于读者对信息系统的理解和学习。

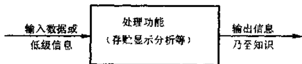


图 1－2　信息系统示意

同信息一样，信息系统也可有层次之分。信息系统的层次，与它所提取信息乃至知识的能力或水平相对应，信息系统层次愈高，所输出信息的知识性就愈强。例如，一种观点认为，从不同管理层次应用的角度，信息系统可分为四类或四种层次。这四种信息系统分别是：第一层，事物处理系统，主要面向具体操作人员的日常活动，负责处理日常事务。第二层是管理信息系统，包含事务处

理系统，提供管理层所需要的数据和信息。再上一层是决策支持系统，它能从管理信息系统中获得信息，具有交互式的分析决策模型，用以支持管理者制定决策。最上层是具有人工智能的系统，即能模仿人工决策处理过程的信息系统。由此可见，这种分类就是信息系统不同层次的一种表现。

5. 数据挖掘和知识发现

最近十几年来，数据采集和存贮等技术，伴随着通讯、互联网等技术的飞速进步，使数据和信息如潮水般地向人们涌来，在世界各地造就了众多的海量、超海量数据库；大量数据来不及处理，闲置在数据库中未能提取信息。为数据海洋包围的人们感觉自己像大海中干渴的水手一样：水手们周围都是水，就是没有自己渴求的淡水；现在人们周围到处都是数据，惟独缺乏自己需要的信息。

在这种形势下，加速发展从数据中提取信息、发现知识的技术的需求和呼声日益高涨。到20世纪90年代中期，这种需求导致形成了一个新兴技术领域：数据挖掘和知识发现（data mining and knowledge discovery，简称DM和KD技术）。数据挖掘和知识发现新技术起源于数据库系统技术领域，它聚焦于在巨型或海量数据库中发现、"认证合理的、新的、潜在可利用的、且最终可理解的结构特征"（Fayyad等，1996），从而提取蕴藏在库中深层次的、新的信息，并形成知识。DM和KD技术形成以来，发展非常迅速，引起人们普遍的重视。

"数据挖掘和知识发现"新兴领域使信息系统与数据库系统技术进一步接近或融合。从信息系统的角度，这一新兴领域的本体思想并不新，因为信息系统之宗旨原本就提取、发掘信息和知识；但是，DM和KD技术是由强大的主流数据库系统技术所驱动的，旨在从海量数据库中系统地、高效地、大规模地提取、发掘信息和知识，这就有不寻常的意义。正如DM和KD技术从各种信息系统技术吸取养分一样，信息系统技术也将进一步得益于DM和KD技术的发展。

信息和知识的关系是一个值得思考的问题。信息和知识都是人类认知过程中，经过去粗取精，去伪存真，由表及里，由此及彼的提炼而获得的，它们都具有客观性、可用性、传输性和共享性。那么，信息是否等于知识呢？按一些国际权威词典，在某些领域中，信息几乎等同于知识。但在信息科学和计算机领域，从多数人的用词习惯来看，信息比知识要低一个层次，介乎于数据和知识之间。例如人们常说，从大量的数据中提取信息，从大量的信息中发现知识。尽管如此，作者认为，如同数据和信息没有绝然的分界线一样，信息和知识之间也没有明确的分界。某些信息其实就是知识；某些人工智能功能较强的高级信息系统所输出的信息，事实上已进入"知识"的领域，或可称之为知识。

§1.2 地理信息系统基本概念和本质特征

本节简明回答地理信息系统"是什么"的问题。

1.2.1 地理信息系统定义

地理信息系统(geographic information systems,简称 GIS),是信息系统的一种,它以表征地球表层空间地理现象和事物的地理空间数据和信息为特定的运作对象。如果用图 1.2 来表示地理信息系统的话,那么,输入、输出两端乃是地理空间数据或信息。换言之,地理信息系统是对地理空间数据进行处理加工、提取有用的地理空间信息乃至知识的系统

不过,在 GIS 领域,地理信息系统定义通常不以这种形式给出。总起来说,GIS 领域的学者习惯于在地理信息系统定义中指明地理空间数据处理的若干主要功能,例如,地理空间数据的采集(collecting)、存储(storing)、显示(displaying)、操作(manipulating)、管理(managing)、分析(analyzing)和建模(modeling,指建立分析模型)等。实际的 GIS 定义中通常列出其中几种功能,但所列不尽相同。此外,特别在中国,GIS 定义中常常指明是在"计算机软硬件的支持下"。例如,一种常见的简明定义是:地理信息系统是在计算机软硬件支持下,对地理空间数据进行采集、存储、显示、管理和分析的技术系统。

有些 GIS 定义详细一些,不仅指明地理空间数据处理的若干主要功能,而且给出其应用目的等内容,以便使初学者和非 GIS 行业的人们更容易了解 GIS 的性质和意义。例如,本书作者在对应用领域的人们宣传时,曾采用如下定义:地理信息系统是在计算机软硬件支持下,对地理空间数据进行采集、输入、存储、操作、分析和建模,以提供对资源、环境及各种区域性研究、规划、管理及决策所需信息的人机系统。

本书不追究 GIS 如何严格定义,主要要求学生把握地理信息系统概念的本质特点。为此,首先需要理解 GIS 运作的对象:地理空间数据和信息。

注意上文中"geographic information systems"的 systems 采取复数形式。事实上,在一般 GIS 英文定义中,以及任何 GIS 作为技术领域总称的场合,systems 均采取复数形式,其意义是很多个别地理信息系统之总和。但当 GIS 指的是单个地理信息系统时,system 就是单数了。在英文文献中,常可以看到"a GIS"的用法。例如,某作者开发了一个用于城市规划的地理信息系统,他就说"developed a GIS for urban planning..."。

1.2.2 地理空间数据和信息

1. 地理空间数据和信息

地理空间数据是符号化表示的、地球表层空间地理现象和事物的直接记录;地理空间信息是指表征地球表层地理现象和事物(或地理环境的要素、物质)的数量、质量、运动状态、分布特征、联系和规律的数字、文字、图像和图形等,以及一切有关知识的总称,是对地理空间数据的解释。应当注意,这里的

"地理现象和事物",包括地球表层空间中所有涉及地理空间位置涵盖的现象和事物。例如,地质、气象、水文等地球科学领域的众多学科,以及土地管理 城市规划等非地球科学领域的学科所涉及的数据和信息,当具有地理空间位置涵盖时,都可以是地理信息系统所处理的地理空间数据和信息。

2. 地理空间数据和信息的三个基本特征

地理空间数据和信息具有三个基本特征 为简明起见,下面仅就"地理空间数据"展开叙述,所述地理空间数据的特征,也适用于地理空间信息。此外,下文中"地理现象和事物"简称为地物。

第一,空间位置特征。地理空间数据必须包括指明地物在地理空间中的位置的成分,这部分数据称为空间特征数据或空间位置数据,有时也简称为空间数据。空间特征数据又有两层含义:其一,地物本身的地理位置(location)。位置通常用某种地理坐标(x,y,z 或经纬度、高程等)或其组合来表达;也可用相对其他参照系或地物的位置来描述,如铁路旁某地物在"铁路以东6m"等。其二,多个地物之间的位置相互关系,或空间关系,如地物之间的距离、相邻、相连、包含关系等。绝大多数空间特征数据具有明显的几何特点,因而有时也称为几何数据。

第二,属性特征。除空间位置以外,地理空间数据还必须包括描述地物的自然或人文属性的定性或定量指标的成分,这部分数据称为属性特征数据或属性数据。例如,表述一个城镇居民点,若仅有位置坐标(x,y),那只是一个几何点,要构成居民点的地理空间数据,还需要其经济(人口、产值等)、社会(就业率等)、资源和环境(污染指数等)等属性数据。

第三,时态特征。时态特征指地理数据采集或地理现象发生的时刻或时段,这部分数据称为时态特征数据或时态数据。同一地物的多时段数据,可以动态地表现该地物的发展变化。时态特征数据可以按时间尺度划分为短期(如地震、洪水、霜冻)、中期(如土地利用、作物估产)、长期(如城市化、水土流失)和超长期(如地壳变动、气候变化)等类型。

综上所述,地理空间数据可以看作是三部分之和,即:

$$\text{地理空间数据} = \text{空间特征数据} + \text{属性数据} + \text{时态数据} \quad (1.1)$$

属性数据和时态数据合起来,描述地理数据的非空间组成部分,有时统称为非空间数据。

时态数据目前应用得还不多。近年来,时态数据越来越受到 GIS 学界的重视;时态空间数据模型(简称时空数据模型)或时态空间数据库新技术,有机地、交互地组织时间维和空间维的数据,推动地理信息系统向动态、多维化方向发展。不过,尽管时空数据模型及技术发展较快,现今实用的地理信息系统一般仍是静态或准静态的:在这样的 GIS 系统中,时态数据通常作为普通属性数据的一部分处理,人们通过属性数据表中的时态数据来调用不同时间的地图 例

如,某地区有一系列历史地图,每幅图的(数据发生)时间放在属性表中;利用这些时间属性数据,可以调用或连续显现各时期地图。由于现在一般的GIS系统还没有达到时态空间数据模型的层次(这种层次目前还主要是研究生的研习内容),本书中,除第七章简单介绍时空数据模型技术的内容外,时态数据皆作为属性数据的一部分对待。如此我们可把式(1.1)简写为:

地理空间数据=空间特征数据+属性数据(含时态数据)　(1.2)

3. 地理空间数据区别于一般计算机数据的本质特征

地理空间数据区别于一般计算机数据的最本质特征,在于下述两点:

第一,空间位置特征。

在上述空间位置、属性和时态三个基本特征中,空间位置特征是地理空间数据区别于一般计算机数据的本质特征。这是因为,一般计算机数据也包括属性和时态特征,只有空间位置特征是地理空间数据特有的;缺乏空间位置特征就不成其为地理空间数据。例如,不少非GIS的信息系统,如户籍管理系统、经济信息系统等,也包括大量街道或城镇居民点不同时间、时段的经济、社会、资源、环境等数据,它们不涉及空间位置和地图表达。如此信息系统的数据就不是地理空间数据。但是,同样是这些数据,若含有表征街道或城镇居民点等地物之地理空间位置特征的内容,成为地理信息系统处理的对象,就成为地理空间数据了。

第二,空间特征数据和属性数据之间的一一对应的关系或链接关系(linkage)。

这一特点在一般GIS书籍没有加以强调,但作者认为,它对学生理解地理信息系统有非常重要的意义。事实上,地理信息系统中每一个地物的属性数据,必须在计算机中挂联到该地物的空间位置数据上,绝不能混淆、错乱。黄河的属性数据,如平均宽度、深度、浑浊度等,不能链接到长江的空间位置上;长江的属性,也不能给予黄河。地理空间数据的这种位置和属性数据间严格的、落实在每一个地物上的一一对应关系,贯穿于整个地理信息系统,制约着一系列地理空间数据的处理功能。例如,点击荧屏上电子地图中某居民点的位置,查询到该居民点的人口、总产值等信息,并显示在地图上该居民点的位置旁,这是GIS最基本的空间查询功能。为了实现这一点功能,居民点的位置数据和属性数据之间就必须有一一对应的挂联关系,使计算机能从一个居民点的位置坐标数据,快速准确地在计算机中找到并取出该居民点的人口、产值等属性数据;反之亦然。

由此,我们可将式(1.2)进一步表述为:

地理空间数据=空间特征数据+属性数据

+两种数据间落实于每个地物的挂联关系　(1.3)

式(1.3)可以说是 GIS 技术独异于、甚至繁难于一般信息系统技术最本质的核心，也是 GIS 技术每一次跟踪计算机主流新技术而变革时，所要进行的技术攻关研究的主要出发点。

4. 有关地理空间数据和信息的用词问题

同很多飞速发展的新兴学科一样，GIS 领域不断冒出各种新名词或新的用词法，以致现在不少用词、用语没有来得及统一或"规范"化，有时甚至较为混乱。为此，有必要就地理空间数据的有关用词情况及本书的做法加以说明。

在 GIS 领域，地理空间数据(geospatial data)，也可称为空间数据(spatial data)或地理数据(geographic data)。在这三个含义相似的名词中，地理数据一词本来是比较适用的，但实际上，习惯用地理数据一词的人不多；相当多的学者喜欢用空间数据一词，也许他们觉得地理数据一词可能误解为"地理学的数据"。

然而，空间数据一词也有问题。首先，空间(space)是一个应用很广泛的名词，在不同学科中含义不同：除物理学理论的 space 外，天文学中的 space 指宇宙空间；行星物理和相关地球物理中的 space 常指地球高层空间和行星际空间；GIS 领域的 space 则是指地理环境或地球表层空间。当代多学科日益交叉，笼统使用"空间数据"一词显然不合适。其次，空间数据一词还被不少学者用来表达地理空间数据的一部分，即空间位置数据(x,y,z 等)部分。在有些 GIS 书籍中，空间数据时而指地理空间数据本身，时而指地理空间数据中的空间位置数据部分，使初学者混淆。

相比之下，地理空间数据一词更为贴切一些。一来它可以避免上述种种混淆；二来与国际上的新专业词"geospatial data"相对应。英文词 geospatial data 既明确，又简洁，因而近年来为愈来愈多的文献所采用。但是可惜，geospatial data 的中译名，无论是"地理空间数据"、"地球空间数据"或"地学空间数据"，都没有原文来得简洁，也不如中文的"空间数据"或"地理数据"简洁。

为此，本书中，凡要求表述较严格的场合，我们将采用地理空间数据一词；但为了简便，同时考虑到国内外采取空间数据(spatial data)一词的学者甚多，我们有时也会采用空间数据一词来代替地理空间数据，例如，空间数据模型、空间数据结构、空间数据分析等。但不论如何，本书将不采用空间数据一词去指地理空间数据中的空间位置数据部分。换言之，本书中，凡是表述空间位置或其相互关系的数据，本书将明确使用空间特征数据、位置数据或几何数据等用词；当使用"空间数据"一词时，表达的是"地理空间数据"本身。

上述关于"地理空间数据"的用词讨论，也适用于"地理空间信息"的用词。

1.2.3 地理信息系统区别于一般信息系统的主要特点

§1.2.2 事实上已给出地理信息系统区别于一般信息系统的根本特点，因为

式(1.1)至(1.3)表明,地理空间数据和信息,与普通数据和信息有明显区别;而地理信息系统以地理空间数据和信息为特定运作对象。本小节在上述根本特点的基础上,进一步分析地理信息系统区别于一般信息系统的几个主要特点。

1. 地理空间数据和信息的特殊复杂性

式(1.1)和(1.3)表明,地理信息系统比一般信息系统复杂。这是因为,地理信息系统的属性数据或信息,是除空间位置及关系以外、所有描述地物自然或人文属性的定性或定量的数据或信息,正相当于一般信息系统所处理的数据和信息。换言之,参照式(1.3),现代 GIS 系统中的属性处理部分,就相当于一个普通的信息系统。由此可见,地理信息系统除了做一般信息系统的工作外,还要额外地运作、处理空间特征数据,及其与属性数据间的一一对应关系;这种处理还进而导致一般信息系统所没有的一系列繁难技术问题。

相应地,地理信息系统的数据库也就比一般信息系统的数据库复杂(参见§1.1.3)。在众多 GIS 系统的数据库中,由于地理空间数据的特殊复杂性,空间特征数据和属性数据通常被分别地加以组织。GIS 的属性数据库部分,相当于一般信息系统的数据库,表现为文字、数字形式的数据集合,通常采用常规数据库软件来管理。但 GIS 数据库技术还必须额外地解决空间特征数据的管理,以及同一地物的空间数据和属性数据协同运作的问题。

2. 必须具备科学可视化功能

可视化功能是地理信息系统的必要条件,而一般信息系统可以没有可视化功能,这是地理信息系统区别于一般信息系统的另一个主要特点(这一点在一般 GIS 书籍中也没有得以强调)。可视化,或科学可视化,已成为现代科学中的热门话题,指通过图形图像等可以看见并认证的手段,来形象表现科学数据的构架和内涵。大量研究表明,可视化能极大地提高信息、知识的理解和传播效率。例如,1998 年美国副总统戈尔在著名的《数字地球》讲演中就指出,单调数字排列形式的数据人们很不容易记忆,但是,科学数据若依据其内部联系排列成可认识的形式,如肖像和星系,那么,极大量的信息转眼间就可为人脑所接收。

地理学是人类科学可视化的先驱领域,人类历史上最早的科学可视化形式就是地图。地图是根据特定的数学法则,将地球上的自然、人文现象,通过制图综合,并以符号和注记缩绘在平面上的图形或图像。这种可视化手段,乃是反映地理事物和现象的时空分布、组合和联系,揭示其发展变化的有效途径。众所周知,关于敌我双方兵力位置的一系列数字罗列对军事指挥员来说,意义是不大的;但是,以地图形式绘出敌我双方兵力部署情况,指挥员就容易分析判断。

地理信息系统利用计算机分析、处理地理空间数据和信息,为资源、环境及其他区域研究、规划、管理和决策服务,自然不能缺少可视化功能。事实上,从地理信息系统诞生之初,"可视化表达地理信息"这一基点就深深植入其设计思

想之中，成为构思地理空间数据模型和结构的出发点之一。正因此，地理信息系统从一开始就与计算机图形图像技术结下不解之缘。

地理信息系统不仅必须具备可视化功能，而且极其重视发展这种功能。经过多年不懈努力，地理信息系统可视化功能从二维到三维，从静态、准静态到动态，从简单到复杂、逼真；同传统地理、地图学相比，GIS 能在高得多的层次上表达地理事物和现象的时空分布、组合和联系，揭示其发展变化。正如我国最著名的 GIS 学者陈述彭先生所指出的，地图学是地理学的第二代语言（第一代是非可视化的语言文字），而地理信息系统是地理学的第三代语言。GIS 是所有研究地物时空分布、组合、联系和发展的学科分支的重要的方法和手段。

3. 区域性和多层次

地理信息系统以地理空间数据和信息为处理对象；而地理空间数据和信息又通常以区域为单位来组织。因此，区域性是地理信息系统的天然特征，不仅纯粹进行区域研究的 GIS 如此，为政府部门和企事业单位管理服务的地理信息系统，也必然面向该单位所管理或涉及的区域。区域沿地球表面展开；地球表面如此广阔，人们通常将地球表面分成很多的图幅来制图，以致区域的分布及特点常需要若干幅水平相接的地图来表达。这一特点导致地理信息系统的数据处理必须具备图幅接边和读图剪切等功能，数据组织管理中需要有图幅管理或图库管理的功能。

地理信息系统还具有鲜明的层次性，而且其层次性包含两种含义。其一，不同比例尺的区域层次。地球上的区域层次是很多的。例如，从我国的小村庄，到乡镇、县（市）、地区、省、大区、国家，直到七大洲。不同的区域层次的地图必须采用不同的比例尺，这个道理在地图学中业已讲过。其二，描述不同地理要素的专题层次，或图层。专题图层相当于地图学中的专题地图，同一区域或同一图幅可以有多种专题图，如杭州市范围的交通旅游图、环境保护图、土地利用图、城市规划图等。在地理信息系统中，不同要素的地理空间数据也常常分别地加以组织，形成同一区域的多重图层或专题数据层次，或用以加强显示的功能和灵活性，或基于它们进行多因子叠合分析。

GIS 的区域性、多层次特点如图 1－3 所示。为简明计，图中仅画出两个区域层次，一个是 1:5 万比例尺的较高区域层次，该层次适宜表达县市级区域范围；另一个是 1:1 万比例尺的较低区域层次，该层次适宜表达乡镇级区域范围。图中可见，1:1 万比例尺层次上的 3 个图幅的范围，对应着 1:5 万层次上一个图幅中的一小部分；当地图从 1:1 万缩编成 1:5 万时，图上地物或细节将过于密集，这就需要进行制图概括（或制图综合）。不同比例尺地图的运作，包括利用计算机进行制图概括，至今仍是 GIS 领域重点研究问题。

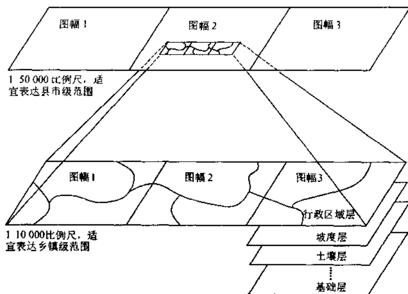


图 1－3　GIS 的区域性、多层次示意

图 1－3 的右下部分还示意出 GIS 系统中另一种层次性，即同一图幅或区域具有多重图层。图中以土地资源适宜性评价为例，进行该评价需要考虑到土壤、土地利用现状、行政边界、坡度等多种因素。当这种评价以乡镇为单位来进行时，就需要利用适合乡镇图幅的土壤、土地利用和坡度等专题图层，以及作为测绘基础的地形图，来对区域内每个具体地块逐一地进行综合分析。在地理信息系统出现前，利用多种专题图进行这种综合分析是很不方便的。而在地理信息系统中，多重专题图层可以直接叠合起来，进行综合分析；分析评价的结果又可形成新的图层。

GIS 中同一图幅或区域分多重图层，也常常是为了增强显示能力和灵活性。为此，有时把同一地理要素的地物也分为若干层。例如，水系分成河流、湖泊等若干层；道路按类型和大小分成若干层；不同大小的城镇也分别分层等。较细的分层使用户打开或关闭图层更加灵活机动；也能在荧屏上缩小或放大显示地图时自动减少或增加图层，以保持地图图面负载的适量性等，起到制图概括的作用。当然，过细的分层也有缺点，这里不一一叙述。

由此可见，区域性与多层次特点也导致地理信息系统比一般信息系统复杂。一般的信息系统也可能涉及区域管理，但是，那只是管理区域的属性数据（有时兼及相关的多媒体浏览）；一般信息系统也可具有层次性，但那里的层次比较单纯，处理技术难度不大。例如，一个工厂的管理系统，厂级、车间、科室或小组等不同层次的数据分析处理，采用常规数据库技术就容易完成。而地理信

息系统的区域性与多层次特点所带来的多重图层、图幅甚至图库管理，以及不同比例尺的地图运作等复杂问题，是一般信息系统所没有的。

4. 数据量较大

同一般信息系统相比，地理信息系统所涉及到的数据量通常要大得多，部分GIS数据库如此巨大，以致人们用"海量"、"超海量"这样的词来描述之。地理信息系统的数据量大，第一来自于它的区域性、多层次特点；第二也是因为地理信息系统包括可视化表达所必须的图形图像数据，而图形图像数据所涉及的数据量经常是很大的。例如，土地管理信息系统包括土地利用现状和规划、土地产权产籍、土地估价等多种数据，其中仅土地利用现状数据，按比例尺1∶10 000计，在浙江省一个中等面积的县级市，就包含十几万个地块图斑和数十万条线（河流、道路、地界等）的数据，而且每条线、每个图斑又挂联着十几、二十几条属性数据，如每个地块的标号、大小、权属、土地利用现状、地块内的细节状况等。

地理信息系统的海量数据，带来了系统运转、数据组织、储存、网络传输等一系列技术问题，这是地理信息系统比一般信息系统复杂的又一个因素。

5. 注重空间分析

地理信息系统同所有的信息系统一样，必须具备分析功能，以提取有用之信息。但是，除了基于数据库技术的一般信息系统分析功能外，地理信息系统还有自己独特的空间分析（spatial analysis）功能。空间分析是地理信息系统中最核心、最重要概念之一。也许正是因为非常重要，这个概念也最缺乏一致明确的界定，众多的学者从不同角度、层面去理解之、定义之、使用之，不同GIS书籍中"空间分析"的涵盖范围多不相同，用词习惯各异。但无论如何，所有的GIS学者都认为，空间分析功能是地理信息系统最重要、最本质的功能。

在本书中，空间分析将从比较广义的角度加以理解，即：空间分析指一切涉及空间位置要素的分析或区域性分析，用以提取地理空间信息乃至关于地物时空分布、组合、联系和发展的知识。按此理解，空间查询检索、空间操作、地图学分析、空间统计分析、空间分析模型等，都属于空间分析的范畴。空间分析模型是空间分析中最重要的一种，不少GIS定义将建模（modeling）功能与分析功能并列。有关空间分析的内容本书还将多次涉及，并将在第五章中集中地讲授，包括对各种空间分析的涵盖及其用词等进行归纳、梳理。

1.2.4 地理信息系统与计算机科学技术的关系

1. 地理信息系统与计算机科学技术总体的关系

地理信息系统是地理、测绘等地学，与计算机和信息科学等之间的交叉领域。总起来说，GIS与计算机技术的关系有三个主要特点：

第一，由于地理空间数据及其处理技术的难度，以及GIS所服务的社会部门管理工作的特殊复杂性，GIS总是如饥似渴地、迅速地吸取计算机主流技术及各种分支技术的各种最新成果，成为计算机科技领域中应用研究内容最丰富的分支之一。正因此，真正的GIS技术专家，必定是计算机技术比较全面的行家。

第二，地理空间数据有不少区别于一般计算机数据的特征，当今计算机和信息科学的一般理论和技术还不能很好解决空间数据结构、组织、处理、表达等方面的问题。因此，地理信息技术有其丰富的内容，它也从特定的角度为计算机和信息科学做贡献。

第三，从历史发展看，GIS与计算机主流和其他分支技术的关系，总的趋势是愈来愈密切，是在"相互融合"。20世纪90年代中期以来，人们常说GIS正在融入计算机主流。事实上，不仅GIS，其他计算机技术，如数据库、互联网、通讯、办公自动化(OA)、多媒体、虚拟现实、CAD和3S技术等，都在不断地连接、融合，形成综合的信息技术。这是现代计算机技术发展的总趋势，这种趋势通常称之为"集成化"(integration)。GIS技术融入计算机主流，就是这一趋势的一部分。

2. 地理信息系统与各计算机具体分支技术的关系

地理信息系统既是地学的一个现代化分支，从计算机科学技术的角度，也可以看作是计算机和信息科学具有地理空间涵盖的分支。从这个意义上说，GIS与各计算机分支技术有着兄弟般的亲密关系。GIS总是积极主动地吸取各种计算机分支技术的最新成果。

在GIS吸取营养的各种计算机技术中，除操作系统和通用型数字文字处理等基本支撑技术而外，与GIS有亲密关系的，首推数据库技术与图形图像处理技术。数据库技术是所有信息系统的主要支撑技术，也是GIS系统的主要支撑技术之一。当然，如§1.2.3.1所述，GIS的数据库比一般数据库结构要复杂。计算机图形和图像处理技术则是GIS"可视化"功能的需要。其中，"图形"用点、线条和多边形(闭合的线条)来表达形象；"图像"则通过画面中所有局部单元的明暗或色调来表达形象。有关GIS图形、图像处理及相应的GIS矢量数据和栅格数据模型，将在第二章详细讨论。总之，数据库技术与图形图像处理技术是地理信息系统与生俱来的、传统的亲缘技术。

网络技术是另一种与GIS关系很密切的计算机技术。由于面向区域分析与管理，GIS天然需要网络化运作，是计算机网络技术最早的应用领域之一。20世纪80年代，发达国家GIS系统就广泛采用当时的网络技术。90年代中期后，GIS领域实现了基于大型关系数据库的、安全高效地运作地理空间数据的网络方式，形成了大型网络级GIS技术。紧接着，GIS又与互联网技术相结合，发展出互联网GIS技术，从而形成了适应现代Internet/Intranet环境的大型网络级GIS技术。

GIS与办公自动化(OA)技术也有天然的渊源。这是因为,办公自动化是网络、文字处理、多媒体、数据库等的集成技术,以提高办公效率为宗旨;而GIS的主要应用领域正是资源、环境等各种区域性规划管理部门,以提高其专业性办公效率、科学化和自动化水平为宗旨。因此,GIS系统天生就需要与办公自动化技术相结合。现在,欲赢得用户和市场的GIS开发者,皆把办公自动化作为GIS系统不可分割的一部分。

此外,人工智能及专家系统等计算机技术,由于其智能化处理数据、提取信息或知识的能力,也一直在GIS领域得到广泛应用。

GIS还与近年来迅速发展起来的其他计算机技术相结合。例如,采用新的计算机组件(COM)开发技术,形成组件式GIS新技术;与虚拟现实(virtual reality或VR)技术相结合,形成虚拟现实地理信息系统(VR-GIS)新技术等;与时态数据库技术相结合,发展时态空间数据模型及技术等。

有必要简单解释一下GIS的图形图像处理与网络运作的关系问题。如前所述,在GIS系统中,表现各种地物的图形图像总是要求挂联属性数据,原则上说。这种挂联关系应当落实到"图元"。在图形方式下,图元指地图中的每个点、每条线和每个多边形(闭合的线条,表示区域),GIS系统要求能查询到每个点状地物(如城镇)、每条线(如河流道路)、每个多边形(如街区)等的属性数据和信息。在图像方式下,图元指每一个像元,要求能查询到地图图像中每个像元的属性数据和信息。从图形图像的图元挂联和查询属性,是GIS最起码的功能或标志,无论GIS系统是否以网络方式运作,皆是如此。

不过,图形图像处理,以及落实到图元的属性数据挂联和查询,增加了GIS网络运作的难度。例如,20世纪80年代末,一种称为客户/服务器(C/S)模式、基于大型关系数据库的安全高效的网络运作技术就已成熟,并开始推广应用。但是,GIS系统实现这种网络运作方式,却是90年代中期后开发出"空间数据库引擎"后的事情。又如,互联网上和多媒体系统中,早就有较成熟的图像运作技术。互联网运作采用浏览器/服务器(B/S)方式,普通用户在网上通过鼠标点击图像也能查询信息。但是,那是以每个图像本身为运作单位(一幅图的局部欲链接信息,该图必须被分割为较小图像的组合),图像中的图元不能单独链接信息。不仅如此,一般互联网和多媒体系统技术不进行点、线和多边形图形运作及其属性数据挂联。而互联网GIS技术,即浏览器/服务器方式的GIS技术,就是要实现网上点、线和多边形图形运作及其属性数据挂联和查询的功能。由此可见,适应现代Internet/Intranet环境的大型网络级GIS技术,包含着很多GIS独特的创新内容。

3. 地理信息系统与一般计算机制图系统的关系和区别

这里所说的计算机制图系统,指的是面向工程设计及工程图绘制的图形处理系统。由于部分行业的工程图采用地图形式,这些制图系统也能绘制精美地图,因而与GIS技术有一定交叉,有必要特别说明。

地理信息系统与一般计算机制图系统的区别,首先在于有无空间分析功能。一般制图系统不挂联地理属性数据,不进行空间查询检索,更不进行空间

分析。其次，地理信息系统的图形处理偏重于地理空间中的自然形态（通常没有解析表达）及其关系；而一般制图系统更擅长规则的、有数学表达的形体的图形处理，因而非常适合于工程设计及工程图绘制。最有代表性的计算机制图系统是 CAD，即计算机辅助设计（computer-aid design）或计算机辅助制图（computer-aid draft）；CAD 技术中最流行的软件是 AutoDesk 公司的 AutoCAD。

如果 CAD 等制图系统仅用来进行与地图无关的工程设计及工程图绘制，那么，GIS 与 CAD 两种技术之间就不会有很多瓜葛。然而，社会上确实存在采用 CAD 技术绘制地图的需求。第一，某些地图测绘工作，如测绘部门的外业地图测绘，不需要空间查询分析，常采用 CAD 制图。第二，社会上有不少行业，如城市规划、水利和电力、电信、雨水、排水、自来水、煤气等管网系统，以及道路等城市公用设施，既需要进行工程设计和施工，又带有区域性行业管理的职能，其工程图大量地采取地图形式，如城市规划图、某种城市管网图等。这些行业的设计施工部门大量地采用 CAD 系统来绘制地图。

然而，CAD 技术缺乏空间分析功能，不适合进行区域管理。因此，上述城市规划、管网、电力电信、水利等行业欲进行区域性行业管理，也需要 GIS 技术。换言之，这些行业具有对 CAD 和 GIS 的双重需求。早先，由于工程设计需要和 CAD 技术普及较早等原因，这些行业曾以 CAD 技术为主；随着 GIS 技术迅速发展和加强区域管理的需求日益紧迫，它们愈来愈多地采用 GIS 技术，形成了 CAD 与 GIS 并用的局面。在我国，这些双重需求的行业内，既有进行区域性行业管理的机构，又有绘制工程图的工程设计单位。前者以采用 GIS 技术为主，后者多采用 CAD 技术。例如，各地的城市规划行业一般都设有规划局、规划设计院等。

注意到上述具有 CAD 和 GIS 双重需求的行业，多是财力和影响较大的行业；在这样的行业中 CAD 和 GIS 技术的相互交织，导致这两种技术的关系引人注目。为此早些年有不少文献专门论述 CAD 与 GIS 技术的区别。不过，现在不同了。近十几年来，地理信息系统与 CAD 等制图系统在技术上不断地接近、融合，这是整个计算机技术领域集成化潮流的一部分。一方面，CAD 技术增设了属性库挂联和其他地理信息系统功能，例如 20 世纪 90 年代中期，AutoCAD 软件专门开发了 GIS 功能的 Map 模块。另一方面，GIS 这些年也大大地加强了图形设计和编辑功能。目前，GIS 和 CAD 两种技术虽然仍有不同侧重和特长，但是，它们之间，以及它们与主流技术之间的融合仍在不断扩展之中。

§1.3 地理信息系统应用和发展

本节重点是论述地理信息系统的应用意义，回答“为什么要学习地理信息系统”的问题。

1.3.1 地理信息系统的社会需求、产生与发展

地理信息系统是一个由社会强烈需求驱动而产生、发展的学科，具有巨大的应用意义和十分广泛的应用领域。

1. 导致地理信息系统产生的社会条件

人类生活在地理环境中，了解地理环境，认识周围事物和现象的时空分布、组合和联系，揭示其发展变化，是人类与生俱来且永无止境的根本需求之一。在古代，对地理空间及其变化的信息和知识统称为"地理"，历史上的地理学涵盖了现代地球科学和环境科学的很多研究领域，如测绘、地质、气象、水文等。如上节所述，认识周围事物和现象的时空分布、组合、联系和变化，需要可视化手段，地图学首先在中国，接着在世界范围内发展起来，成为地理学的第二代语言，使地理学成为人类科学可视化的先驱领域。地理学和地图学关于地理环境的知识，被广泛应用于政治、军事、经济等领域。

从文艺复兴时代开始，接踵而来的世界范围的地理大发现和地理制图技术的革新，促进了近代地理学的诞生。随着地理学各分支或地理环境中各圈层研究的深人和知识积累，地理学中的测绘、地质、气象、水文等分支相继成熟，具备了形成独立学科的条件；这种分化是人类从各个专门角度全面、深化地研究和认识地理环境与地球的标志。

进入20世纪以后，人类发展的步伐明显加速，资源、环境等社会问题亦日趋尖锐，要求提高资源、环境及各种区域性管理的范围、力度和科学化水平的呼声日趋高涨。与此同时，由于生产力及科学技术突飞猛进，人类对地球表层空间及资源环境的数据采集、分析和处理的能力迅速提高。特别是：第一，20世纪中前期，立体航空摄影和遥感成像技术的发展，使摄影测量工作者能以很高的精度快速地进行大面积测量，为地理学家、土壤学家、生态学家、地质学家、土地利用专家等提供了前所未有的观测数据。第二，地图学及测绘制图技术的长足进步，又为地学家们提供了前所未有的存储、显示、分析与处理地理空间信息和数据的功能。由于具备了"需要"与"可能"两方面条件，人类对地球表层环境的认识迅速扩展和深化，形成了包括地理、地质、地球物理、地球化学、大气科学、水文、水力、灾害等众多分支在内的大学科门类——地球科学，兴起了有多种分支的环境科学和一系列开发、利用和管理各类资源的分支学科，建立起土地管理、城市规划、旅游管理等一系列进行区域性规划、管理和决策的政府机构或部门；兴起了各种相应的经济领域或产业。

然而，随着地学、资源、环境和区域科学，以及相应的社会经济领域的快速发展，传统的手工处理方式愈来愈不能适应形势发展的需要。例如，以手工方式制图，数据存储和查询能力很有限，且难以实施地图的更新、增补、定量分析

等;特别是区域性规划管理和决策,需要进行多因子空间综合分析,而每一个因子通常表现为一幅专题地图。例如,进行图1-3所示的,某区域土地资源适宜性评价,不难想象,在人工方式下,对照分析土壤、土地利用、坡度状等多幅专题地图,来评价每一块具体的土地,是非常困难的。为此,在GIS技术产生前,分析者曾用透明薄膜绘制区域的各项专题图,并将它们对齐、叠置在灯光桌(玻璃桌面下放置电灯)上,逐一观察地图的每个局部并进行分析评价。

形势的发展,要求对地理空间数据采集、存储、处理和分析的方式和技术手段进行变革。如果说20世纪中期对这种变革的需求日益紧迫的话,那么,这时期前后新兴的科学技术,特别是计算机技术,以及现代系统论、信息论、控制论等,为这一变革提供了最重要的条件和保证。

2. 地理信息系统产生和发展

地理信息系统产生于20世纪60年代北美地区的两股源流。第一股出自土地资源分析管理的需要。60年代早期,加拿大国家土地调查局为了分析、处理大量的土地资源调查资料,开始建立地理信息系统;经过十多年的试验与研究,于70年代初投入产品生产。该系统一开始就面向空间分析(基于栅格数据形式的空间叠置分析)。如果以空间分析作为地理信息系统的必要标志,那么,该系统无疑是60年代出现的世界上第一个地理信息系统。第二股源流出自计算机地图自动绘制的需要。美国哈佛大学的计算机图形与空间分析实验室,于60年代开发出世界上第一个通用的制图软件包CYMAP。这是世界上开发基于图形方式(矢量数据格式)的通用型地理信息系统的成功起点。

地理信息系统一诞生,就不断迅速地吸取计算机主流技术的每一个最新成果,随着计算机和信息技术的飞速进步而发展,可以说,计算机和信息科学发展有多快,地理信息系统或地理信息科学的发展就有多快。60年代中后期,新出现的地理信息系统很快受到相关领域的高度重视,美国或北美建立起若干个地理信息系统组织,如城市与区域地理信息系统联合会(URISA)和美国航测与遥感协会(ASPRS)等。70年代,地理信息系统迅速发展与巩固。美国等发达国家投入了大量的人力、物力、财力,开发出众多各种特色的地理信息系统,商业GIS软件开始活跃;不少大学开始培养GIS人才,开展GIS学术活动。

80年代是地理信息系统大发展、GIS普及和推广应用阶段,地理信息系统由比较简单的、单一功能的、分散的系统发展成为多功能的、用户共享的综合性信息系统,并向智能化发展。80年代后期,两种主要的GIS技术(栅格和矢量)实现一体化,相应地,GIS技术与遥感技术一体化,成为新一代(第二代)GIS技术成熟的标志。若干优秀的地理信息系统软件,如美国环境系统研究所(ESRI)开发的Arc/Info等,在此期间被广泛推广应用。GIS软件性能的提高,促使地理信息系统应用迅速扩展:应用的领域从传统的资源和环境,迅速扩展到众多领

域;应用的社会部门从政府扩展到涉及区域规划管理的企业和商用领域;应用的地区扩展到全世界。GIS市场空前扩大,形成一个巨大的产业。GIS教育在发达国家高校普及,学术活动空前活跃。

90年代是地理信息系统发生飞跃的时期,其标志是:第一,地理信息系统发展形成了一个世界范围公认的、包括基础理论研究体系和技术应用体系在内的完整学科:地理信息科学或地球信息科学。第二,90年代中后期,在加速发展的互联网/企业网(Internet Intranet)技术和面向对象开发技术等,以及整个计算机技术领域集成化趋势的带动下,GIS技术发生飞跃,形成第二代GIS技术。GIS与多种计算机技术相集成,共同融入计算机主流技术。技术的进步促使地理信息共享和社会化程度极大提高,相应地,GIS应用领域和GIS产业的发展更加迅猛,GIS走向公众,进入日常生活领域。GIS教育进入中学,GIS学术活动空前深化。1998年,美国副总统戈尔提出"数字地球"科技发展战略。数字地球是"可嵌入超海量地理相关数据(Geo-referenced data)的,我们行星多重比例尺的三维的表达"。数字地球的提出,进一步提高了GIS学科和技术的地位,为地理信息科学和技术的发展提供了一个强大的新推动力。

3. 地理信息系统在我国的发展

地理信息系统在我国亦称为资源与环境信息系统。我国地理信息系统的研制与应用,20世纪70年代末才开始,但发展很快,大体可分为四个阶段。第一,70年代准备阶段。在这一时期,蓬勃开展的遥感技术应用研究为该阶段后期我国正式组建GIS研究准备了条件。从1978年到1979年,我国进行舆论准备,正式提出倡议,开始组建队伍,组织个别实验研究。

第二阶段从1980年到1985年,为起步阶段。中国GIS正式起步,以1980年中国科学院遥感应用研究所成立全国第一个地理信息系统研究室为标志,在这几年的起步发展阶段中,我国地理信息系统在理论探索、硬件配制、软件研制、规范制订、区域试验研究、局部系统建立、初步应用试验、技术队伍培养等方面都取得了进步,积累了经验,为在全国范围内展开地理信息系统的研究和应用奠定了基础。

地理信息系统进入发展阶段的标志是从第七个五年计划(1986—1990)开始,地理信息系统研究作为政府行为,正式列入国家科技攻关计划,开始了有计划有组织有目标的科学研究、应用实验和工程建设工作。众多的应用领域开展了地理信息系统研究与开发工作,完成了一批综合性、区域性(全国的或地方的)和专题性的信息系统(或数据库)的建设,开发了一系列地理信息系统基础软件或专题应用软件;开始进行地理信息系统教育培训,出版有关地理信息系统理论、技术和应用方面的著作,并积极开展国际合作,参与全球性地理信息系统的讨论和实验,全国范围内形成了地理信息系统的科研队伍,逐步建立了不

同层次、不同规模的研究中心和实验室,特别是1985年和1990年分别成立了对我国GIS事业做出重要贡献的,中国科学院资源与环境信息系统国家重点实验室和武汉测绘科技大学测绘遥感信息工程国家重点实验室。通过近五年的努力,我国在地理信息系统技术上的应用开创了新的局面,并在全国性应用、区域管理、规划和决策中取得了实际的效益。

自90年代起,我国地理信息系统步入快速发展阶段。实施了地理信息系统和遥感联合科技攻关计划,"遥感、地理信息系统和全球定位系统的综合应用"被列入国家"九五计划"重中之重的科技攻关项目。强调地理信息系统的实用化、集成化和工程化,力图使地理信息系统从初步发展时期的研究实验、局部应用走向实用化和产业化,为国民经济重大问题提供分析和决策依据,并在一批经济相对发达、技术力量比较雄厚、用户需求急迫的地区和城市或部门建立了成功的GIS实用系统。成功地建立数字化测绘技术体系,建设国家基础地理信息数据库和基础环境数据库,发展了基于地理信息系统的数据产品研制与生产。推进国产GIS软件系统,特别是基础通用软件系统的研制及其实用化,国产GIS软件与国外的差距迅速缩小,涌现出若干有市场竞争力的地理信息系统软件。在重大自然灾害监测与评估系统的建设及应用、地理信息系统规范化与标准化研究等方面也成绩斐然。经过十年的发展,中国地理信息系统事业取得了重大的进展,向高水平、产业化方向前进。

1.3.2 地理信息系统的应用领域

地理信息系统的应用领域非常广泛,GIS专业人才的就业形势乐观,近几十年来,直至现今,国内外GIS专业人才供不应求的态势从未改变。

1. 85% ~90%的政府部门

政府是国家的行政机关,政府及其各部门从全局或从某项业务角度对所辖区域实施行政管理;而地理信息系统通过对地球表层人文经济、自然资源及环境等各种地理空间数据和信息的管理和分析,把握城乡和区域的自然环境和社会经济要素的空间分布、空间结构、空间联系和空间过程的演变规律,能为政府进行区域性决策,如资源和环境管理、区域发展规划、公共设施管理、交通安全等方面的决策提供依据。因此,地理信息系统有时被称为"政府技术";政府及其部门是地理信息系统的天然用户,在任何国家地理信息系统事业起步时,政府及其部门总是第一批用户。

国内外大量的调研表明,85% ~90%的政府部门需要应用GIS技术。

表1-1给出我国当前应用GIS技术较多的管理部门,该表是作者根据自己对GIS实践了解整理出的,难免有漏误之处。但是,读者就此已可以感受到众多政府部门对GIS的需求。

表 1-1 我国当前应用 GIS 技术的领域或部门

<table>
<tr><th>应用领域</th><th colspan="2">内 容</th><th>所涉及的管理部门或机构*</th></tr>
<tr><td>区域规划</td><td colspan="2">区域总体规划</td><td>计划委员会</td></tr>
<tr><td>土地资源和资产管理</td><td colspan="2">(较大范围区域)国土资源整治
农村土地管理
城镇土地管理</td><td>国土办
土地管理机构,农业局,农业区划办
土地管理机构,城市规划局,房地产局</td></tr>
<tr><td>水利与水资源管理</td><td colspan="2">水旱灾害
水工程管理,水资源管理,农田水利和小流域管理,流域规划</td><td>防汛防旱指挥部和水利部门的水文局等机构
水利部门的水工程管理、水政管理、水文局、农田水利和水利规划院等机构,环保局</td></tr>
<tr><td>旅游管理</td><td colspan="2">旅游资源、设施等管理</td><td>旅游管理局</td></tr>
<tr><td>林业管理</td><td colspan="2">森林资源、设施等管理</td><td>林业局</td></tr>
<tr><td>水产管理</td><td colspan="2">水产资源管理</td><td>水产局</td></tr>
<tr><td>地质矿产</td><td colspan="2">地质矿产资源管理
石油资源管理</td><td>国土资源部门的地质矿产管理机构
石油管理部门</td></tr>
<tr><td>人口管理</td><td colspan="2">人口统计分析
计划生育</td><td>人口普查办
计生委</td></tr>
<tr><td>环境管理</td><td colspan="2">环境保护
环境监测
生态和生物多样性等</td><td>环境保护局
环保局所属环境监测站
环保局所属自然资源管理部门,农业、林业局等</td></tr>
<tr><td>公安和急救</td><td colspan="2">犯罪空间分析
110 报警系统
119 报警系统
灾害评估</td><td>公安局
公安局
消防局
保险公司,民政部门</td></tr>
<tr><td rowspan="2">城市管理</td><td colspan="2">城市规划及管理
房地产管理
园林管理
城市综合整治</td><td>城市规划局和设计院
房地产局,土地管理机构
园林管理局
市政府或城市综合整治部门</td></tr>
<tr><td>城市管网管理</td><td>排水(雨水 污水)系统和道路维护
城市给水系统管理
煤气管道系统管理
电力线路网(地上、地下)管理
电信线路网(地上、地下)管理</td><td>市政工程管理局
自来水公司
煤气公司
电力局或电力公司
电信局或电信公司</td></tr>
<tr><td rowspan="2">交通运输</td><td>城市交通</td><td>交通、道路规划设计
城市交通管理
城市公共交通</td><td>城市规划局
公安局交警或巡警大队
公交公司</td></tr>
<tr><td colspan="2">公路运输和航运管理
航空运输
铁路运输设计和管理</td><td>交通局及其设计院
民航总局
铁道部或铁路运输管理、设计部门</td></tr>
<tr><td>卫生</td><td colspan="2">健康和疾病防治</td><td>卫生局、防疫站</td></tr>
<tr><td>农业病虫害</td><td colspan="2">农业病虫害防治</td><td>农业局</td></tr>
<tr><td>军事</td><td colspan="2">众多内容(略)</td><td>军事部门</td></tr>
</table>

* 因各地机构不尽相同,且因时而变,部分管理机构名称姑妄称之。

2 企业和商用领域

GIS的传统应用领域是环境和自然资源评价、规划、交通、公用设施管理和军事等，以政府或准政府部门应用为主。在美国等发达国家，80年代GIS应用就向企业、商用领域迅速扩展，到90年代，GIS在企业、商用领域的广泛应用导致形成了一个新领域——business geographics。1992年，以这个新词为名的全美专业杂志创刊。business geographics不是传统意义上的"商业地理"，其含义是利用GIS技术，特别是空间分析和应用模型技术（有时也称"地理技术"，即geographic technology），去做business。例如，进行选址、选线路、市场分析、系统分支网点的布点分析、精细农业分析等。这个领域增长得如此之快，以致世界著名的GIS学术期刊之一，美国航测与遥感协会会刊PE&RS在1996年底总结GIS十年间发展时，说这个领域"也许是90年代GIS应用的最大增长领域"。近几年来，经济全球化加速，公司集并愈来愈大，对GIS空间分析的需求继续增长。如果说，众多政府部门业已为GIS施展身手提供了广阔领域的话，那么，GIS进入企业和商用领域，为GIS带来了更大的应用空间。

3. 民用

随着新技术的发展，能够获取或共享的区域性社会经济等数据越来越丰富，美国等发达国家在90年代不仅将GIS应用向企业、商用扩展，而且使GIS走向公众，走向民用，进入日常生活领域。例如，基于GIS和全球定位系统的民用轿车导航系统已经很普及，旅行者开车到哪儿，导航系统荧屏上就给出那里的比例尺合适的电子地图。

综上所述，GIS可应用于涉及区域分析和管理的各种政府部门、企事业单位和民用领域，应用领域非常广阔；广泛的应用又会导致教育和研究部门中GIS工作岗位的增加。因此，GIS人才的潜在需求量是相当大的。在美国等发达国家，GIS人才的就业形势一直很好。

在中国，GIS应用总的特点是：第一，发展很快，应用领域和范围高速度扩展。第二，应用规模同发达国家相比仍有较大差距。后者主要表现在，一方面，我国的GIS应用还局限于部分政府部门的部分业务，尚未进入企业、商用和民用；另一方面，我国社会发展水平和管理水平还不够高，空间分析技术或地理技术还远未能得到应有的应用。即令如此，GIS人才目前在我国仍比较缺乏，很多政府部门，甚至待遇较好的中小城市土地管理部门，都找不到所需的GIS人才。今后，如果我国的GIS应用能像发达国家那样，向更多的政府部门及其业务，进而向企业、商用乃至民用领域扩展，我国的GIS人才需求还将有相当大的潜在空间。

GIS的应用领域究竟有多广？一言以蔽之：凡是需要使用地图的领域，都是地理信息系统的应用领域。

1.3.3 地理信息系统与地理信息科学

1. 地理信息科学的形成

地理信息系统作为一种可以在地球科学、有关资源与环境学科及区域管理学科领域通用的技术，以一种全新的方式去组织、认识、理解和使用地理空间数据和信息；地理信息系统的应用改变了地理信息分发和交换的方式，这种新方式天然地有利于多学科之间的沟通和交叉研究。地理信息系统在自身的研究发展过程中，以及在它应用于多学科领域的实践与研究过程中，必然会逐渐形成一门以处理地理空间数据和信息的理论、技术为研究对象的，对地球科学、有关资源与环境学科及区域管理学科领域有普遍意义的学科。

经过30余年发展，到20世纪90年代，地理信息系统及遥感等领域已经积累了从理论、技术到应用研究的非常丰富的内容。包括我国在内的各国学者皆认为，GIS事实上已形成一个新的学科领域；这个学科在美、英等国称为地理信息科学（geographic information science）；在另一些国家为geomatics（或类似的词），我国通常译为"地球信息科学"。地理信息科学一词是著名GIS学者Goodchild于1992年提出的。1995年，美国地理信息科学大学联合会（University Consortium for Geographic Information Science）成立，并于1996年提出地理信息科学的十个重点研究领域，即空间数据获取和集成、（网络）分布式计算、地理表达的扩展、地理信息的认知、地理信息的互操作性、比例尺问题、GIS环境下的空间分析、空间数据基础设施的未来、地理数据和GIS运作中的不确定性，以及GIS和社会。1997年，英国出版的《国际地理信息系统期刊》（International Journal of Geographical Information Systems），更名为《国际地理信息科学期刊》（International Journal of Geographical Information Science）。1999年，美国测绘学会的刊物《地图学和地理信息系统》（Cartography and Geographic Information Systems），更名为《地图学和地理信息科学》（Cartography and Geographic Information Science）。同年，在我国，由GIS和遥感领域首席学者陈述彭院士主编的本领域期刊《地球信息》，也更名为《地球信息科学》。凡此种种，都表明地理信息科学或地球信息科学已被公认形成为一个重要的新兴交叉学科。

2. 地理信息科学的涵盖

同地理信息系统一样，地理信息科学的涵盖也远远超出地理科学本身。现在一般认为，地理信息科学是地理、测绘等地球科学与计算机、信息科学之间的交叉学科。从一方面看，它是地球科学或地理学的信息化、现代化分支；从另一方面看，它可以说是计算机和信息科学具有地理空间涵盖的分支。由于是新兴学科，地理信息科学还没有一致公认的定义。本书将地理信息科学定义为：研究地理空间数据的采集、存储、表达、处理、分析、建模和地球信息机理的交叉学

科。

地理信息科学具有自己的基础理论研究和技术应用体系。在基础理论研究方面，美国国家科学基金会与美国地理信息科学大学联合会1999年举行的专门研讨会纪要文献《地理信息科学：一个新兴交叉研究领域中的若干紧迫问题》认为，地理信息科学基础研究的重点课题包括：软件和数据集成、比例尺和分辨率、处理模型、可利用性（人机交互、用户界面设计等）、地理信息的表达、不确定性和认知等。该联合会又于2000年将数据挖掘和知识发现、遥感数据和信息、地理可视化、GIS本体论基础和分析地图学研究补充列为地理信息科学的重要基础研究课题。

地理信息科学的技术应用体系，首先是3S技术。3S技术指：

地理信息系统技术（GIS），遥感技术（remote sensing，简称RS）——采集、接收遥感，特别是卫星遥感数据，并从中分析、提取地球资源环境各种信息的技术，以及全球定位系统技术（global positioning systems，简称GPS）——利用系列卫星实时高精度确定地面目标精确位置的技术。

20世纪90年代早期，GIS、RS与GPS集成，形成一体化的3S技术。3S技术有时亦笼统地称为GIS技术，因为3S技术的核心是GIS，RS与GPS采集和处理的数据，最终要集成到地理信息系统中进一步处理和分析；再者，现代GIS技术的内涵十分丰富，业已覆盖一部分RS与GPS的内容。例如，GIS新技术的一种，即"移动GIS"，就是结合嵌入GPS和通讯技术的GIS技术。

除3S技术外，地理信息科学的技术应用体系还包括3S技术一般不能涵盖的现代测绘技术和地球信息分析技术等。

地理信息科学是支持"数字地球"科技发展战略的核心的深层次的学科领域。

3 有关新名词讨论

地理信息科学导致出现了若干专业新名词或名词新用法。

首先，新学科的出现使GIS一词具有更宽广的含义，因为地理信息科学（geographic information science）英文的直接简写，也是GIS。因此，现在GIS可以从两个层面理解：从狭义上，GIS指地理信息系统技术，即面向区域、资源、环境的规划、管理和分析，处理地理空间数据的计算机技术；广义上，GIS可以指地理信息科学，即一个具有理论和技术体系的学科。当同时论述GIS技术和GIS学科时，为避免混淆，美英一些文献用GIS指技术，而用GIScience或GISci来简写地理信息科学。

其次，如前所述，除"地理信息科学"外，新学科还可称为"地球信息科学"等。地球科学，简称地学，英文为Earth Science或Geo-science，但地球信息科学的外文通常采用Geomatics或类似的组合简写词。应当说，"地球信息科学"一

词更利于体现新学科在自然科学中的定位，因为新学科是包括地理、测绘和其他有关地球科学分支，与计算机、信息科学之间的交叉学科。

不过，本书仍采用"地理信息科学"。这是因为，第一，本书及其作者属于地理科学范畴。第二，"地理信息科学"是美、英等国的用法，而它们在本领域最前沿的学术地位使这个词通行世界的可能性较大，就像"地理信息系统"一词现在通行世界那样。第三，地理信息科学一词在学术上也是合理的，因为在地球科学领域内，别的一级学科皆从某个局部角度来研究地球表层环境，只有地理科学是从全局的角度综合研究地球表层环境；第四，GISci不能覆盖地球科学与计算机、信息科学交叉领域的全部（如地球科学中很多数理方程的数值计算研究不可能归入GIS的范畴，参见§2.1.2）。

4. 本书的定位

既然从整个地理信息科学的视野观察，"地理信息系统"一词主要指技术，那么，本书名为《地理信息系统》，是否主要是一本技术讨论书呢？不然。本书并不展开技术方面的深入讨论，主要学习地理信息科学的一些基本概念和原理。因此，本书实际上是一本地理信息科学的基础的入门的书籍，书名为《地理信息科学基础（或初步）》更为确切。

既然如此，为什么要称为"地理信息系统"呢？这是因为，它是几年前国家确定的课程名称。在我国，国家确定的学位点、专业和课程的名称不可能经常地变更，以致难以跟上一些发展特别快的学科名称变化的步伐。本领域的专业和学位点（硕士点、博士点）名称也是如此。按照当今学科发展的实际情况，本领域学位点和专业名称现在最好都改称为"地理（或地球）信息科学"。但是，本领域学位点名称至今仍称为"地图学和地理信息系统"；本科专业名称仍称为"地理信息系统"。

§1.4　地理信息系统的功能及组成

本节比前两节更进一步，具体地介绍地理信息系统的功能及组成，以便读者概略了解地理信息系统的主要框架。

1.4.1 地理信息系统的主要任务

前面说过，地理信息系统是对地理空间数据进行处理加工，提取有用的地理空间信息乃至知识的系统。这里，"提取的地理空间信息乃至知识"，可以具体一步地归纳为五个方面的内容。这也是地理信息系统所要回答的五个基本问题。

1. 某个地方有什么

这一项基本任务，指利用地物或区域的空间位置数据或信息，来查询、获取

地物或区域的自然或人文的有关属性信息。在地理信息系统中，地物或区域及其位置可以有多种表达方式，如地图图元或图形、地理坐标、地名或邮政编码等。因此，"某个地方有什么"的任务，就是允许用户能通过表征地物或区域位置的地图图元或图形、地理坐标、地名或邮政编码等，查找和得知这些目标或区域有什么特征、内容或地物组成成分等，例如，让用户了解某区域内各种地块的所有者、土地利用现状和估价等信息。GIS 回答用户的方式也是多种多样的，可以采取文字、列表方式，但更多是以可视化方式，例如在荧屏或输出地图上，以鲜异的色彩或高亮度显示满足指定条件的所有特征。

2. 符合某些条件的实体在哪里

"符合某些条件的实体在哪里"的问题，即利用地物或区域的属性数据或信息，来查找满足给定条件的地物或区域的地理位置信息。为此，地理信息系统应当允许用户能通过多种方式，例如在荧屏上从预定义的可选项中选取、填写表格或逻辑表达式等途径，来指定一个或一组条件；然后回答满足用户指定条件的地物或区域在哪里的问题，并加以显示。例如，某地区内，距离某种污染源 200m 以外、满足若干地质条件的地下水源最可能在哪里？

3. 变化趋势

变化趋势(trends)指已发生和正在发生的变化。地理信息系统要求能根据已有的数据，包括现状的和历史的数据，识别地理现象已经发生或正在发生的变化，总结出变化趋势，为用户提供关于变化趋势的图表等。例如，某一区域中有多少土地利用状况发生了变化？其中，有多少柑橘地转为他用？分别转为何种用途？这种变化可回溯到多少年？历史上不同时段的变化趋势有何差异？增加了还是减少了？各时段长度如何？等等。

4. 空间结构特征

空间结构特征(patterns)揭示已有和现有的地理实体或现象之间的空间关系。地理信息系统应当允许用户能按指定的区域组合地理空间数据，探究与地理事件有关的其他地理空间因素，分析其间的 patterns 或空间相关关系，并采用各种方式向用户报告分析结果。例如，某社会学家利用地理信息系统研究城市犯罪，发现某城市暴力刑事案件与酒吧分布之间的相关关系：暴力事件较频繁地发生于距离任一酒吧 100m 范围内，当地时间下午 5 点至夜里 2 点之间。这就是一种 pattern。不仅如此，他还可利用地理信息系统进一步研究这种"酒吧暴力事件"与该城市的照明情况、交通格局和警察巡逻方式的相互关系，从中得出有意义的结果。

Pattern 一词常用于某些现代科学研究中，但难以找到简单而确切的中译名，有时译为模式、样式、格局、结构特征等。这里译为"空间结构特征"，是想尽可能接近 pattern 一词用于这里所具有的含义。一般认为，从研究数据中发现 patterns，是研究和认识过程的重要一环。

我们在§1.1.3中曾提到关于"数据挖掘和知识发现"的较流行定义（Fayyad等，1996）。该定义认为，"数据挖掘和知识发现"就是the process of identifying valid, novel, potentially useful and ultimately understandable patterns in data，我们译为"在数据中认证合理的、新的、潜在可利用且最终可理解的结构特征的过程。"

5 模拟

模拟（simulation）比发现Patterns更进一步，即解决某个地方如果具备某种条件会发生什么的问题。地理信息系统要在发现变化趋势和空间结构特征的基础上，进行更高层次的研究，建立分析模型并加以验证，发现具有普遍意义的规律性，建立新的数据关系，产生解决方案。例如，上述社会学家可以在研究某城市"酒吧暴力事件"与当地照明情况、交通格局和警察巡逻方式的关系的基础上，进一步利用地理信息系统和现代网络环境，对其他城市进行类似研究；一旦发现带有普遍意义的规律性，就可以把研究推向更高层次：建立通用的分析模型，进行面向未来的预测和决策。

1.4.2 地理信息系统的主要功能

地理信息系统的功能服务于地理信息系统的上述任务，覆盖了从数据采集、处理和分析，乃至决策、应用的全部过程。尽管目前GIS软件和系统千差万别，它们的基本功能大体相同。Maguire等按照地理信息系统中的数据流程，将地理信息系统的基本功能分为五类（10种）：①采集、检验与编辑；②格式化、转换、概括；③存储与组织；④分析；以及⑤显示。其中，分析功能（包括空间分析与模型）被称为地理信息系统的高级功能。在上述功能中，数据存储、检索和人机交互显示，贯穿于整个流程。下面，我们就按Maguire等基于数据流程的思路，来分述地理信息系统的主要功能，其中少数分类及用词我们按自己的理解做了调整。

1. 地理信息系统的基本功能

• 数据采集和输入。数据采集和输入指地理信息系统从现实世界的观测，以及从现存文件、地图中获取地理空间数据，并输入到计算机中。目前可用于地理信息系统数据采集和输入的方法、技术很多。一部分是通用的计算机方法，如利用键盘、软驱、光盘机、扫描仪等来输入数据；另一部分是地理信息系统及相关领域的专业方法，如利用全站仪等测量设备采集地理空间数据，利用手扶跟踪数字化仪从原有地图上采集并输入地理空间数据等。

• 数据处理。主要包括检验与编辑、格式化、转换和概括。地理空间数据输入到地理信息系统中之后，需要进行检验，并按照地理信息系统数据库运作和分析功能等的要求进行编辑、修改，以保证数据在内容与空间上的完整性、数值逻辑一致性、美观、正确性、适宜于用户的具体需求等。这就是检验与编辑功

能。

格式化(formalization)功能是使 GIS 中的数据采取,或转换为某种适合地理信息系统运作和用户需要的数据结构,例如,矢量数据结构或栅格数据结构,或其相互转换等

数据转换通常有两种类型。一类是坐标转换。坐标转换又包括两种:一是不同平面坐标间的坐标变换,变换公式可以是线性的,也可以是非线性的。例如,人们常常从 GIS 输入设备(数字化仪或扫描仪等)所给定的英寸单位的坐标系,向我国通用的高斯–克吕格大地坐标系转换;二是地图投影,即将地表曲面(球面或旋转椭球面)上的地理坐标,转换为地图上的平面坐标。

另一类数据转换是不同数据格式(format)的转换。这种转换是由于 GIS 软件开发商和数据生产者皆采用各自的方式来组织地理空间数据,导致实际应用的 GIS 数据格式千差万别。例如,美国的 ESRI(环境系统研究所,Environmental System Research Institute,Inc.)、MapInfo 和 Intergraph 等公司所开发的 GIS 软件包,所用数据格式皆不同;即使是同一公司的不同软件产品,如 ESRI 公司的 Arc/Info 和 ArcView,也会采用不同的数据格式。另一方面,GIS 数据生产者或供应商,如美国地质测量局、农业部、人口调查局等,所提供的地理空间数据的格式也各不相同。而地理信息系统用户常常需要采用多种 GIS 软件开发商或数据供应商的数据。因此,地理信息系统必须具备进行不同数据格式之间转换的功能。

数据概括功能对应着地图学的制图概括(generalization,也称制图综合),即利用计算机手段来进行制图概括。地图学讲过,制图概括很大程度地依赖于地图学家的专业智能及人为处理。因此,实现数据概括功能的难度较大。目前常见的 GIS 数据概括功能仅有:线条的简化、平滑;部分实现比例尺缩小时地物数量、形态等的筛选和简化表达(参见图 1–3);以及在荧屏上缩小或放大地图时,自动地减少或增加图层,以保持适当的地图图面负载量等。总的来说,GIS 系统的数据概括功能至今仍很欠缺,需要尽可能采用计算机人工智能技术、加强这方面研究开发。

• 数据存储、组织和管理。为了实现 GIS 数据处理等功能,地理空间数据必须按 GIS 表达和运作的要求来组织和管理,而这种组织与管理又与地理空间数据的存储技术密切相关。在一般计算机系统中,数据的存储、组织和管理,主要是数据结构和数据库系统技术的任务。但是,由于地理空间数据的特殊复杂性(空间特征、属性特征及挂联关系),一般的计算机数据结构和数据库系统技术还不能很好地解决 GIS 的数据存储、组织和管理的问题。因此,地理信息系统必须发展自己特有的数据存储、组织和管理功能。

GIS 的数据存储、组织和管理功能不像其他功能那样,能够较简单地加以解

释（详见第二、三两章的讲解），因此，现在进一步讲述没有意义。目前只要理解一点就行了：地理信息系统的数据存储、组织和管理功能，是[illegible]GIS [illegible]达、运作和其他功能的基础。

● 显示与输出。可视化是地理信息系统的基本特征。除采用报告、表格和图表等常规数据表达方式外，所有的地理信息系统皆为用户提供可视化表达地理空间数据的手段或输出产品。例如，通过计算机荧屏输出或硬拷贝（纸张等）输出等形式，显示、分析、绘制和输出地图及其相关数据。地理信息系统十分强调良好的、交互式的制图环境和可视化分析环境，让用户能够设计和制作出高质量的地图，或采用人机交互方式进行空间分析研究。

● 空间查询与分析。空间分析是地理信息系统的核心功能，也是 GIS 与其他计算机系统的根本区别之一。现在多数人认为，空间分析包括空间查询，空间查询是地理信息系统空间分析功能中，最基本的低层次分析功能；简单空间查询以外的空间分析，特别是模型分析功能，属地理信息系统的高级分析功能。由于空间分析的重要性，下面专门加以论述。

2. 空间分析与模型分析功能

空间分析在地理学，乃至其他地学研究中有着悠久的历史与传统，数学概念与方法的引入，从统计方法扩展到运筹学、拓扑学等方法的应用，加强了其定量分析的能力。而地理信息系统的空间分析功能，使传统的空间分析能力大大加强，能更科学、高效地分析和解释地理特征间的相互关系及空间模式。

地理信息系统的空间分析可分为三个不同的层次。

● 空间查询检索。空间查询检索是地理信息系统最基本的分析功能，能否有效地从地理信息系统海量数据库中检索出所需信息，将影响地理信息系统的进一步分析能力。空间查询检索一般有三种形式：一是从地物的空间位置查询或检索地物的属性；二是反过来，从地物的属性数据条件查询或检索地物的空间位置特征；三是让用户通过图形手段，查询指定范围中地物的空间位置和属性双重特征。将查询检索的结果用图形图像等可视化方式表达出来，也是空间查询检索功能的必要组成部分。

● 定式化的空间分析功能。定式化的空间分析功能是本书作者首次使用的用语，因为作者认为，很有必要将这一类空间分析划分出来。空间分析的种类很多，在大量的空间分析研究和实践过程中，一些常用的对不少应用领域都有普遍意义的空间分析手段被总结出来，被做进通用型 GIS 商业软件包中，成为其中定式化的空间分析功能模块，如空间叠置（或叠加）分析模块、网络分析模块、三维分析模块、空间聚类分析等。这里以前者为例加以说明。

空间叠置分析（overlay）是利用多重专题（属性）图层的，具有普遍应用意义的区域性多因子综合分析，图 1-3 右下部分所示的土地资源适宜性评价，就是

种空间叠置分析。由于普遍应用，空间叠加分析业已成为某些通用型 GIS 商业软件包中的定式化模块。叠置分析模块允许通过人工干预，按自己应用对象的具体要求来进行叠加分析，并为用户可视化地表达分析结果，非常方便。例如，研究生物多样性保护的用户，可利用叠置分析模块对地形、土壤类型、降水量、植被、土地利用、地籍（土地权属）等属性要素（或专题图层），进行灵活的叠加分析，一目了然地观察分析结构，评价不同的区域开发规划对保护生物多样性的影响。

• 其他的空间模型分析。更多的空间分析不具有跨领域的普遍性，它们仅面向自己的应用领域，通常需要结合本领域，如环境、水利和生物生态等的专业模型或专家系统。我们可以笼统地称之为空间模型分析或应用模型分析。应用模型分析中，有的是将专业模型（如水文－水力模型）整块地植入 GIS 系统中；有的则是将应用领域的知识与 GIS 功能有机交融。还有一种特别值得一提的空间模型分析，是地理学理论上空间分析。

空间分析的内容很多，我们将在第五章详细讲授。

我们平日所说的“GIS 软件的分析功能”指上述前两种，即空间查询检索功能和定式化的空间分析功能。各种 GIS 软件分析功能的强弱程度和涵盖范围不尽相同。美国 ESRI 公司的 Arc/Info 被认为是具有丰富空间分析功能的 GIS 软件的代表。在不少情况下，人们只需要应用一些较简单的空间分析功能，为此，一批制图功能较强，但空间分析功能不太强的 GIS 软件得以发展。这些软件现在常称为桌面绘图软件（desktop mapping）。从英文可见，这种翻译不确切，因为这里的“绘图”，不是一般绘图，实际上是“绘制地图”。目前，有代表性的桌面绘图软件有 MapInfo 公司的 MapInfo 软件、ESRI 公司的 ArcView 和 AutoDesk 公司的 AutoCAD Map 等。MapInfo 公司曾宣称自己的软件不是 GIS 软件。不过，从地理信息科学领域全面分析，桌面绘图软件都至少具有空间查询检索等空间分析功能，因而都是 GIS 软件，只是 GIS 分析功能不够强罢了。

1.4.3 地理信息系统的主要组成成分

同一般信息系统一样，地理信息系统也包括数据及信息、硬件、软件和人员（开发或应用人员）四大要素，只是这里的“数据及信息”，是地理空间数据和信息。

1. 地理空间数据和信息

地理空间数据和信息的概念已在 1.2.2 小节中详细讨论，这里不再赘述。地理空间数据和信息，采取图形、图像、文字、表格和数字等各种形式，是地理信息系统运作的对象，是 GIS 所表达的现实世界经过模型抽象的实质性内容。

2. 硬件系统

硬件系统是计算机系统中的实际物理装置的总称，是 GIS 的物理外壳。硬

件系统可分为计算机主机、各种输入输出外部设备、网络传输设备等主要部分，其间的物理装置，可以是电子的、电的、磁的、机械的、光的元件或装置。地理信息系统的规模、精度、速度、功能、形式和使用方法，乃至软件，都与硬件有极大的关系，受硬件指标的支持或制约。反过来，所有的硬件组件协同工作，完成用户通过软件系统所下达的各种指令。

3　软件系统

软件系统是指挥硬件运作的各种必需的程序，地理信息系统的档次和质量，主要由 GIS 软件来体现。地理信息系统的软件系统由系统软件和 GIS 专业软件组成。系统软件指各种最基础的软件，GIS 专业软件建立在系统软件的基础上。专业软件又包括两个层次。一是面向 GIS 用户群的通用型商用软件或软件包，二是针对特定用户的应用分析程序。应用分析程序通常是地理信息系统的用户或开发人员，根据其特定的管理应用或分析模型编制的，用于某种特定应用任务的程序，这是用户最为关心的，真正用于分析研究的部分。

地理信息系统的硬、软件系统还将在下一小节中进一步介绍。

4．系统开发、管理和使用人员

尽管现代计算机和 GIS 技术已相当先进，人始终是人机系统中的最活跃最重要的因素。

人的作用体现在两个方面。一方面，人是地理信息系统的设计者与建设者。这样的人通常称为系统开发人员。系统开发的工作不仅限于系统建设期间，还延伸到系统建成之后，例如在应用管理人员参与下的系统扩充、完善和更新等。

另一方面，人又是地信息系统的应用者或服务的对象。这样的人称为用户。用户不仅要应用 GIS 系统进行区域性分析管理，灵活利用空间分析功能提取多种信息，为研究和决策服务，而且必须进行系统的完善、管理、维护、数据更新等工作。用户与系统开发人员之间必须密切合作。

地理信息系统不仅依赖于开发、管理和使用人员，还与社会化的"人"及相关因素，即社会组织、体制等因素密切相关。地理信息系统的社会化特点比一般信息系统明显。

地理信息系统的应用重点是区域分析管理及相应的社会部门，特别是各种政府部门，因而比一般的信息系统更多地受到社会环境因素的影响和制约。这个特点在我国更加显著。有些学者认为 GIS 组成成分应包括社会因素，认为 GIS 应当有 5 个，甚至 6 个成分，以强调社会因素在 GIS 中的作用。例如，美国威斯康星大学学者曾给出一个包含六个构成成分的 GIS 定义：地理信息系统是硬件、软件、数据、人、组织和习惯体制所构成的，对空间数据进行采集、输入、存储、操作、分析和建模的系统（GIS are systems of hardware, software, data, people, organization and institutional arrangements for collecting, storing, displaying, manipulating, analy

zing and modeling spatial data)

1.4.4 地理信息系统的软硬件组成

与一般的信息系统或计算机系统相比，地理信息系统的软硬件组成有其特点，而且通常要复杂些，有必要专门讲述。

1. GIS 硬件系统的特点

同一般的计算机系统一样，GIS 硬件系统的基本组件也包括执行程序的中央处理器（CPU）、存储器、外部存储介质、输入输出设备、网络设备等。但由于图形图像可视化运作和海量数据特点，地理信息系统的硬件系统有两个比较明显的特征。第一，它对硬件的性能要求，例如对运算速度、内存和硬盘大小等的要求比较高；第二，对输入输出设备的图形图像运作性能的要求也比较高。地理信息系统的某些常用的输入输出设备，是一般计算机系统很少应用的。由于这两个原因，GIS 硬件系统相对“昂贵”一些。

图 1－4 是一幅 GIS 软硬件系统构成的示意图。注意：该图着重展示地理信息系统的软、硬件构成特点，而忽略一般计算机软硬件共性的内容。例如，该图不展示CPU、存储器等计算机具体硬件，而较多地描述GIS所偏重的输入输

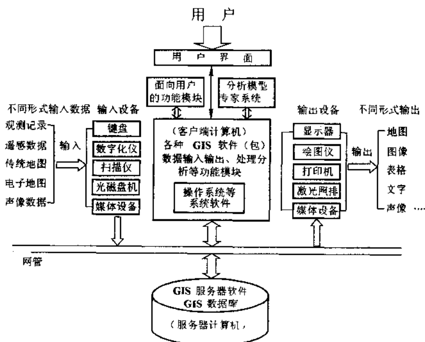


图 1－4　地理信息系统软硬件组成示意图

出设备。考虑到当代 GIS 皆是网络级的，该图给出网络型的 GIS，各部分通过网络连接，但是，该图并不具体描述网络模式或结构。事实上，若在图 1-4 中去掉网管，把输入、输出和数据库都直接连到中央方块上，那就是一个单机版的 GIS 软硬件构成示意图。

2. 输入、输出设备

图 1-4 的左边为地理信息系统的输入设备或输入端。输入端的最左边列出了地理数据的几种主要形式，五个小方框给出地理信息系统的常用输入设备。首先，“观测记录”指通过手工或仪器直接在野外或其他地点所采集的数据，这些数据常采取文字或表格记录形式。其中，手工数据常通过计算机键盘输入到计算机中，数字化观测数据（用电子测量仪器测得）则可通过磁盘、软盘等输入到计算机中。“遥感数据”指人造卫星或其他飞行器对地观测的照片或电子数据，前者可通过扫描仪，后者可直接通过光磁盘机等输入计算机。观测记录和遥感数据是比较原始的 GIS 输入数据。

输入端的地图则是过去的信息产品，在这里作为初级信息被输入到地理信息系统中，成为进一步处理加工的素材。其中，传统的“纸张地图”是地理信息系统最常见的输入材料，它们通过数字化仪或扫描仪，被转换为计算机能理解的数字形式而进入计算机（参见 §2.2.4）。地理信息系统所用的扫描仪常是大型或较大型的扫描仪，同数字化仪一样皆是一般人少用的外部设备。

“电子地图”指业已数字化的地图或数字地图，可通过光磁盘机等直接输入计算机。现代地理信息系统已愈来愈多地采用商业化的电子地图产品作为输入数据，例如美国地质测量局的地形图和美国人口调查局的“老虎”（TIGER）数据，就是有众多用户、价廉物美的电子地图产品。近年来我国也开始有了电子地图产品，但商业化程度较低。

图 1-4 的右边五个小方框列出地理信息系统的常用输出设备，以及地理信息的多种输出形式。输出设备中，除显示器即计算机荧屏外，绘图仪和较大型的打印机也是地理信息系统常用，而一般人少用的输出设备。又由于地图常要出版，高性能的彩色激光照排机也是某些地理信息系统的常用输出设备。

此外，由于多媒体技术的发展和普及，多媒体设备和声像形式的地理数据或信息，也愈来愈多地成为地理信息系统的输入输出设备和输入输出形式。

3. GIS 软件系统

图 1-4 的中间部分是地理信息系统的计算机主机部分，我们将 GIS 主要软件功能都画在这里。在现代网络方式下，GIS 数据和数据管理的功能主要放在服务器一端，GIS 系统从服务器端检索、调用数据，在客户端进行多种 GIS 处理。目前读者无须追究“客户/服务器”的概念，请将主要注意力放在图 1-4 的中间方框图中的内容上。

图中可见，系统软件是GIS系统最核心最基础的一部分，包括操作系统、面向各行各业的通用数据库系统管理软件和一些通用型应用软件(如微软office软件等)。GIS专业软件建立在操作系统、通用数据库系统管理软件的基础上，并可调用所需的通用型应用软件的功能。GIS专业软件又包括两个层次。一个是面向各种GIS用户群的通用型GIS商业基础软件，包括GIS软件包、计算机图形软件或计算机图像处理系统等。这些软件系统通常又按各种GIS基本功能构建，形成地理信息系统的各种模块或子系统，例如，地理空间数据的输入子系统、数据存储与检索子系统、GIS数据库管理子系统、数据处理与分析子系统、输出子系统等。

GIS基础软件包还包括近年来出现的互联网GIS(Web GIS)软件和空间数据库引擎软件(也称中间件软件等)，它们分别为互联网环境下和"客户/服务器"网络环境下的地理空间数据运作服务。当前，客户/服务器网络模式基于大型商业关系数据库技术。换言之，在现代网络级GIS系统中，位于服务端(参见图1-4中下部)的大型商业关系数据库系统，不仅能帮助管理地理空间数据，还能进行网络管理。

本节提到的一些GIS新技术名词，读者只须初步了解，不必细究。这不会影响继续学习。

GIS软件的另一个层次是针对特定用户的应用分析程序，它们又建立在通用型GIS商用软件之上，是地理信息系统的最上层建筑。

4. 地理信息系统中的"上层建筑"

GIS"上层建筑"如图1-4中上部所示，主要由面向用户的功能模块，以及分析模型、专家系统等组成，它们是用户最为关心的，真正用于分析研究的部分。该处还绘有"用户界面"(user interface)的方框，这是用户与GIS系统(或任何计算机系统)直接进行交互的界面，它也可以看作是地理信息系统"上层建筑"的一部分。

为什么需要这些上层建筑呢？这是因为，地理信息系统发展了数十年，虽然从底层开发GIS基础软件的努力至今未断，但现代绝大多数地理信息系统直接采用商业化GIS软件产品。商业化GIS软件产品为了赢利、发展，总是面向尽可能多的用户群体，因而主要开发有普遍应用意义的GIS功能模块。但现实中的GIS系统总是为具体用户服务的，GIS用户除需要通用功能外，必定还有其特殊的需求。为了满足这种需求，GIS商业软件通常都为用户提供进一步开发的手段。由此缘故，具体的GIS应用通常基于用户选用的商业GIS软件平台；用户利用商业软件平台提供的开发工具或开发语言，或者利用更加通用的计算机软件工具或语言，如微软公司的VB、VC等，来调用商业平台的功能或组件，组装或开发针对自己特定需求的功能模块、分析模型或专家系统，以及相应的

用户界面。此项工作常称为"二次开发"，有时也称为"定制"，就像利用商业平台提供的材料"搭积木"，或为顾客量体定制衣服一样。

图1-5是一个现代视窗(Windows)风格的GIS水利应用系统的用户界面(部分)，其中左半部为系统一级主菜单，右半部显示二级菜单(部分)。用户利用鼠标点取、激活任一菜单项，即可打开该项功能。图中，此刻主菜单中被激活的是"灾害分析和评估"项。因此，右边出现的是"灾害分析和评估"的二级菜单。

从主菜单可以看到，该系统包括几个面向用户的功能模块。一是面向防汛指挥部的"汛情显示分析"模块，该模块可用多种方式形象地分析、显示省内各报汛站、雨量站传来的水、雨、汛情实时数据。二是"流域规划"模块，该模块为流域规划及治理管理的部门服务。三是"水工程、工情管理"模块，该模块为水工程管理的科学化、自动化提供辅助决策手段。此外，该水利系统是一个基于大型关系数据库的客户/服务器模式的大型网络级GIS系统，因此，该系统还为水利用户的系统管理员准备了一个"后台维护"模块。每一个功能模块都有很丰富的内容。

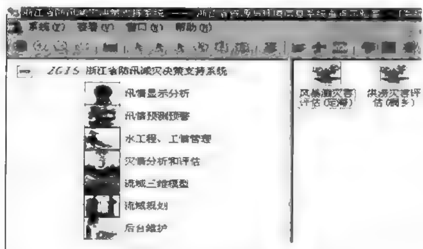


图1-5 一个水利应用系统主界面(部分)

除了上述面向用户的功能模块以外，主菜单中还有"汛情预测预警"、"灾害分析和评估"和"流域三维模型"几项，这几项都是面向用户的GIS应用模型。例如，汛情预测预警模型乃是GIS与水文-水力学模型之结合，基于流域三维地形(含河流断面)数据，计算和预测一定雨情条件下的水、汛情，并将预测的结果通过GIS"可视化"功能，直观、形象地表达出来。

图1-6是采用另一种常见风格的一个城镇土地产权产籍管理信息系统的

界面。该界面的上部为系统的主菜单，这里没有应用模型，给出的全部是面向城镇土地管理业务所需要的11个功能模块。每一个功能模块都有很丰富的内容。为了进一步方便用户，系统还制作了一系列图标（分别位于界面中部和右部），让用户能便捷地调用最常用的功能。例如，一旦用户点取右边最上面标有"图形"两字的方图标，系统将立即跳出专门进行图形编辑的界面和工具。图中的中右部还有一个大的"浏览产权详情"的对话框。对话框也是一种界面形式。该对话框后的背景中还可看到一幅城镇地籍图。这说明图1-6的界面不是系统的初始界面，而是系统查询、处理某一块土地时（此刻是权利人为"张志伟"的土地）的过程界面。

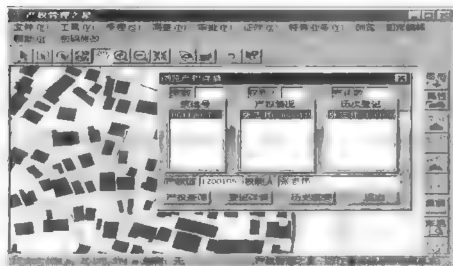


图1-6 一个地籍管理系统的用户界面（部分）

通过上述两个例子，不难理解用户界面与用户的功能模块和分析模型之间的相互关系。用户界面是用户借以操纵、使用计算机系统的媒介，现代计算机的用户通过用户界面（界面上的菜单、图标等）来调用所需要使用的功能或分析、应用模型。因此，用户界面必须简明、一目了然，使用户易懂、易用。

第二章
空间数据组织与计算机表达

从本章起，我们将从各方面具体地探讨地理信息系统。本章首先讨论、GIS怎样在计算机里组织地理空间数据，以表达地理空间事物和现象。这是学习后面章节的基础。

下一章，即第三章将讲述与本章内容密切相关的空间数据管理问题。在本章与下章中，读者将会较多地遇到“空间数据结构”、“空间数据模型”等可能不大熟悉的专业用语；这些用语在每一本GIS书中都会用到。考虑到本书将面向很多的自学者，他们也会参阅其他GIS书籍，我们在本章第一节中将穿插地编写一些其他GIS书籍不一定给出的背景知识。

§2.1 从地理空间现象到计算机世界

2.1.1 从现实世界到计算机世界

本小节先从计算机科学的一般性角度，简介利用计算机模拟现实世界中事物或现象的一般概念，重点了解“数据模型”和“数据结构”等词的原意。应当说明，这些概念即使在众多计算机科学书籍中也缺乏一致公认的定义。因此，下面的段落基于本书作者自己的理解而编写，主要目的是帮助读者理解GIS书籍常用的“空间数据模型”和“空间数据结构”等概念。

1. 模型与概念模型

客观世界无限丰富多彩，任何研究都必须将所关注的局部世界加以简化和抽象。在现代科学技术中，对现实世界的抽象和简化表达，通常称为模型。地图就是一个借助于符号和注记，形象表现所关心的区域内容的模型。建立模型的过程称为建模(modeling)。模型的详尽程度不尽相同。通常把对现实世界的第一层简化和抽象，称为概念模型。概念模型给出所研究的主要事物的概念及其相互联系的框架。在计算机应用以前的自然科学中，人们在概念模型的基础上，建立数学模型，进行基于解析表达式或数学方程的计算。例如，应用各种简

化的流体力学方程计算大气和水体的运动等。

计算机兴起后，不仅很快地解决了数学方程的数值计算问题，而且迅速将研究的触角伸展到没有解析表达式或列不出数学方程的各种非数值计算领域，成为愈来愈强有力的 能用多种手段，包括可视化等多媒体手段来支持科学研究的工具 在计算机中，现实世界是以数字和各种符号形式来表达和记录的，并被最终表示为 进制形式来进行操作 因此，基于计算机的科学研究，需要将经过人脑简化、抽象的研究对象用数据来加以描述，并将数据组织为计算机所理解、支持的形式。这个从现实世界到计算机世界的转化过程，可以大体分为概念模型、数据模型、数据结构和文件格式四个主要层次。

2. 数据模型

数据模型(data model)的概念主要来自于计算机科学的数据库系统理论 数据库是为现实世界的 -部分(用户单位)服务的。为此，数据库技术首先借助概念模型，把握用户世界中的实体及其相互联系的概念框架。概念模型按平常人的思维方式建立，以方便用户理解和交流，不依赖于具体的计算机系统。为了具体地定义和操作数据库中数据，指导计算机运作，还要将概念模型转化为数据模型(图2-1)。按照著名数据库专家E. F. Codd的理论，数据模型实质上是一组为用户服务的规则，这些规则规定其数据结构如何组织，以及应当允许进行何种操作。

数据模型通常由三要素组成：数据结构、数据操作和完整性约束条件。其中最基本的要素是数据结构。数据模型通常以它所基于的数据结构名称来命名，如基于网状数据结构的数据模型称为网状模型等。

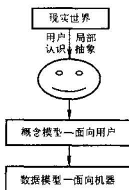


图2-1 现实世界至计算机世界

3. 数据结构

数据结构是计算机科学中与数据库系统关系较密切的一个基础性的分支。它主要研究没有数学方程的非数值计算领域，如表格、树状、网状、图结构等的程序设计问题中，计算机的操作对象以及它们之间的关系和操作。数据结构具体指同一类数据元素中各元素之间的相互关系，包括三个组成成分：数据的逻辑结构，数据的存储结构和数据的运算。

数据的逻辑结构是对数据之间逻辑关系的描述，有时也简称为数据结构；数据的存储结构又称为物理结构，是数据的逻辑结构在计算机存储器中的实现；数据的运算是在数据的逻辑结构上定义的、物理结构上实现的操作算法。

例如，数据间的逻辑结构可能是线性结构(结点间是一对一的关系)、树形

结构(一对多关系)和网状结构(多对多关系)等,可采用顺序、链接、索引或散列等存储方法,并定义和实施检索、插入、删除、更新、排序等操作算法。

由此可见,数据结构涉及到计算机运作中比较基础的层次。上述数据库理论中的数据结构,大体相当于本小节数据结构中的逻辑结构。

总的说来,数据结构和数据库系统做的都是计算机数据的组织工作。但是,数据结构的工作更"基层"一些,且不涉及具体的用户;而数据库系统涉及较大数据集合的组织和管理,其应用与具体用户有关,因而涉及到模拟、表达客观世界局部的问题。数据结构有时就称为数据组织,而数据库系统到达"数据管理"的层次。

4. 文件格式

数据结构处理数据元素关系和操作的共性。在相同的数据结构下,不同软件厂商或数据生产者,还会采取不同的具体形式来组织、存储数据,即采取不同的文件格式(参照§1.4.2.1)。例如,图像结构的数据就有 TIF、BMP 和 JEG 等数十种文件格式。文件格式涉及到更具体、实际的数据组织问题,这也是地理信息系统在组织和运作地理空间数据时必须处理好的一个环节。

2.1.2 GIS 中地理空间数据组织的主要对象

从地理空间现象或事物到计算机世界,一般也有概念模型、数据模型、数据结构和文件格式几个层次。这个过程有时统称地理空间数据建模。本小节先主要从概念模型的层次上,讨论地理信息科学中,地理空间数据建模的研究对象和表达对象。注意,地理空间是人类生活的空间,现有的大量学科,包括地球科学各分支和地球科学以外的其他学科,绝大部分的研究对象实际上都是地球表层空间中的事物和现象,因此,笼统地说地理信息科学的数据建模对象是地表空间现象或事物,是不够确切的。地理信息科学只能从自己特定的视角对地理空间现象或事物进行数据建模和研究。

1 数值计算领域的地理空间数据组织

我们认为,地理信息科学的数据建模,主要是对非数值计算领域而言的事实上,同一般数据一样,地理空间数据的组织和处理,在数值计算和非数值计算领域中也有明显不同。在地球科学中的一些领域,人们抽象出用方程表达的数学模型,设计解方程的计算机算法,编出程序、求得结果,从而得知物理量怎样因空间位置而异,怎样随时间而变化:

$$\text{求解的物理量(即属性)} = f(x,y,z,t)$$

这里,(x,y,z) 表示地理空间位置,t 表示时间。在数值计算领域,地理空间数据依据方程的要求而组织,结构相对简单

地球科学中需要进行数值计算的,主要有两类方程 一类是数理统计方

程,另一类是数学物理方程。众所周知,数理统计方程在大量学科中广泛应用,例如地学中的地理、大气、地质、海洋、水文等领域都有众多的统计模型。在计算机应用以前,这些方程有时无法解决(没有解析解);即使能解,解方程的过程也常常十分繁难。回归方程计算就是一个熟知的例子。

在大气科学、海洋科学、水力学、测绘科学、地球物理学等学科中被广泛应用的数学物理方程(常常是微分方程组),一般比数理统计方程更复杂,数据量、计算量更大,而且很多求不出通解,因此更需要应用计算机。以大气科学为例,较理想状态下的大气运动可以用5个流体力学或热力学的基本方程来概括,只要解出这个方程组,就可以较好地预测给定起始条件和边界条件下的大气运动。但是,这5个方程没有解析表达式的通解。计算机出现以后,气象学家才得以借助于计算机巨大的运算能力,再加上某些简化方程的条件,用差分代替微分(用 Δx 代替 dx 等)的方法,来求解大气运动方程,从而预测天气甚至气候。这就是现在所说的“数值预报”方法。在现代大气科学研究中,基于数学物理方程的数值方法已成为主要手段,数理统计方法已降到辅助地位。其他地球科学分支也有类似的情形。

在地学实际研究工作中,还常常既采用数学物理方程,又采用数理统计方程,借助于计算机,各尽其用,达到解决问题的目的。利用计算机求解方程的数值计算,在很多涉及地理空间数据运作的地球科学和其他学科中,发挥着重要的作用。

2. 再谈地理信息科学之涵盖

从字面上看,地球(或地理)信息科学应当涵盖地学方程的计算机数值计算,因为地学方程都是在研究地球表层或地理空间中的物质流和信息流。然而实际上并非如此。地球科学中绝大多数基于数学物理方程及数理统计方程的计算机数值计算,是在物理学和数学理论的基础上,大气、测绘、地质、水利等相关学科理论与计算机数值计算方法相结合的结果,从学科涵盖上,这些研究不可能归入 GIS 的范畴,它们既不是 GIS 的主要关注对象,也不是 GIS 的特色和成果。

可见,除了不多的地理数理统计方法外,GIS 中所要组织和处理的地理空间数据,主要是非数值计算领域的地理空间数据。计算机科学中的“数据结构”分支所关注的表格结构、树状结构、网络结构、图结构等,事实上都是 GIS 抽象表达非数值计算领域的地理事物和现象所关注的对象。当然,GIS 的涵盖还要大得多。

上述讨论实际上牵涉到地理科学的涵盖问题。我们常说,地理科学是研究地理事物和现象的时空分布、组合和联系,揭示其发展变化的学科,这种说法其实不够确切。大气、水文、海洋、地质等学科由于能较好地反映大气或水体等的时空分布、组合和发展变化,相继从古老的地理学独立出去;而仍然以整个地理环境为己任的地理科学,至今没有基础的数学物理方程,除了数理统计描述外,仍然

主要只能借助于定性分析和研究。地理信息科学为解决这个问题带来希望。

非数值计算领域的地理空间数据组织和运作的研究具有非常重要的意义。这是因为，第一，地球表层物质和运动极其复杂，大量的事物和现象不能用数学物理方程描述。因此，非数值计算领域的地理空间数据研究和运作具有极其广阔的空间。第二，GIS 着重非数值计算领域的地理空间数据的研究，并不等于不应用数学物理方程和数理统计方程；相反，GIS 经常与应用领域的数理方程相结合，并且以自己强有力的空间数据库和可视化等功能支持这些数学模型。例如，GIS 与水力/水文模型相结合，不仅可以用地形数据库支持水力/水文模型的输入，而且可以生动形象地表达水力/水文模型的结果，如洪峰演进、淹没区模拟等。

3. GIS 所抽象、表达的地理事物和现象

GIS 怎样组织数据以模拟地理事物和现象呢？地球表层物质和运动极其复杂，不难理解，为了用计算机研究没有方程表达的地理事物和现象的时空分布、组合和发展变化，地理信息科学首先必须把所关注的地物抽取出来，并加以简化、抽象，建立概念模型，以便计算机表达各种地理事物或实体的空间位置、形状和属性，以及它们之间的空间相互关系和其他相互联系。

例如，我们关注一条河边有一个人和一只足球。注意如前所述，地理空间是人类之居所，地理空间的事物几千年来已被人类从很多学科的角度来加以认识。在这条河边，艺术家可能因这只足球的形态而具体地描绘它；物理学家可能从力学角度计算这只足球因受力加速运动将到达的位置；水力学家则可能利用流体力学理论来计算水体运动；还可能有若干专家从多个学科角度来研究这个人，等等。但是，现在 GIS 是把它们作为地理事物来进行自然科学的考察，关注的是此时人、足球与河的空间位置及其相互关系。因此，我们需要将人、足球与河分别简化，抽象为一个点和一条(或两条)线，表达出三者的相互位置关系，并用一一对应的属性数据去说明这两个点和这条线是人、足球与河，而不是其他的什么东西(图 2-2)；有必要的话，还可用进一步的属性数据去说明这只足球是这个人踢到河边现有位置的。

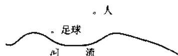


图 2-2 空间对象的抽象和简化

我们将 GIS 所抽象、表达的地理事物和现象，称为空间对象；空间对象的位置相互关系，称为空间关系。以上例可见，为了表达空间对象及其空间关系，我们还必须用几何手段去描述空间对象本身的形状及其所占的空间位置。同时，用属性数据去说明空间对象及它们之间的空间关系和其他相互联系。这就是 GIS 数据建模的主要对象。

鉴于空间对象和空间关系概念和重要性，下面进一步讨论之

2.1.3 空间对象和空间关系

1. 空间对象

地理空间中的事物和现象，有的是连续分布于整个研究区域，如大气场、温度场、地球重力场、地表高程等。一部分连续分布的物理量通过数学物理方程求解，不是 GIS 的研究对象，但 GIS 可以用可视化手段去表达之，支持之。地理空间中更多的事物和现象只占据空间的一个局部，仅在该局部内连续，具有比较明确的边界。这是 GIS 所主要模拟的空间对象。

一部分空间对象本身的大小在研究中可以忽略，因而可以用一个点来表示；另一些具有不可忽略的空间延展性，它们需要用线、面、体或体表面来表示。这里，体或体表面指三维的空间形体，体具有不等于零的体积；而面在 GIS 中，请特别注意，一般指平面地图上之区域（英文常用 area），或者围成区域的闭合曲线，即多边形（polygon）。之所以称为多边形，是因为地理信息系统中的曲线，实际上是由很多足够小的线段组成的（荧屏上放大时，曲线的局部都表现为折线），看上去的闭合曲线实质上是很多小线段首尾相连形成的多边形。

总之，空间对象可以采取零维（点）、一维（线）、二维（面）和三维（体或二维曲面）等不同空间维数的几何形体来表达。从这个角度，空间对象可分为点状空间对象、线状空间对象、面状空间对象和体状空间对象。

除空间维数特性外，空间对象还可以从其复杂性、规则性、人为性等角度认识或区分。有些空间对象比较简单或单纯；另一些则可采取复合形式，即复杂空间对象可包含若干简单空间对象。例如，一个湖泊可包含若干小湖泊和岛屿；杭州西湖就包含三个岛，以及被苏堤、白堤分隔开来的三个子湖泊。空间对象有规则和不规则之分，规则的空间对象可以采用解析表达，例如可用直线、矩形或圆等图形来表达。空间对象还有自然的和人为的之分，不少人为的（man-made）空间对象及其边界在地球表层并没有明显可见的标记，例如行政区边界（如省界、县界等）和人类依据各种因素对地球表面所作的各种分类分区（如农业或气候区划）的边界等。

有些复杂自然对象如河流、珊瑚礁、地震、洪水、农田、矿体等，用常规的欧氏几何很难表达。但许多自然对象具有相似的形体属性，可以用分形几何来描述。现实世界中的许多空间对象具有分形空间维，许多自然表面的分形维在二维与三维之间。

空间对象是对地理空间实体简化、抽象的结果，而这种简化、抽象属于人类认识的范畴。因此，不同的研究者对同样的空间实体进行简化、抽象的结果，即空间对象，可能不同或不尽相同。例如，采样时的条件限制差异，可能导致抽象

出的空间对象不同。一个典型的例子是,城镇边界、河流等线状空间对象的形状,会因沿边界的采样点的疏密程度而不同。空间对象也可能因研究者给予的定义差异而不同,例如,某植被覆盖区的大小和范围可能随着定义该植被覆盖的阈值差异变化而变化。

上述空间对象的各种特征和差异,都是构架空间数据模型时需要考虑的问题。

2. 空间关系

这里说的空间关系主要从狭义上而言,指几何上空间对象之相互关系(广义上,空间关系有更多的内涵)。空间关系的描述是多种多样的:有定量,也有定性的;有精确,也有模糊的。各种空间关系的描述具有一定的联系。空间关系的认识不仅与空间数据模型有着密切联系,并且与人类对空间对象的感知、语言的表达和心理因素有着密切的联系。

空间关系通常可以分为三类:度量空间关系、顺序空间关系和拓扑空间关系。

- 度量空间关系。度量空间关系(metric spatial relationship)主要指空间对象之间的距离关系。地理空间通常被看作是欧氏空间,欧氏空间中的距离即一般观念的真实距离。欧氏距离有直线距离(如飞机航线)和沿线路而行的旅行距离等之分。

空间距离的概念还被地理学家加以扩展。例如,"时间距离",以空间对象之间旅行所需的时间来表达二者距离;"经济距离",以空间对象之间某种经济关联的程度来表达二者距离;"风险距离"以空间对象之间旅行的安全程度来表达二者距离;还有"社会距离"、"认知距离"、"生态距离"等。

- 顺序空间关系。顺序空间关系(order spatial relationship)描述空间实体之间在空间上的排列次序,这种顺序总是对特定的前提而言的。经常涉及的顺序空间关系有三类。一是上下顺序关系,它基于地球重力方向,可以借空间对象的绝对或相对高程来体现。二是前后顺序关系,该关系是相对于某个研究对象而言的:距离其近者为"前",距离其远者为"后"。第三类是基于东南西北地理方向的顺序空间关系,例如空间对象 A 在空间对象 B 之东或东北,A 较 B 偏西等,这种顺序关系可利用地理坐标体现之。

点状空间对象之间的顺序空间关系比较容易研究;但涉及线状、面状或体状空间对象的顺序关系时,情况可能复杂起来,这主要是因为它们本身的形状可能较复杂,以致难以确切地定义其顺序关系。

- 拓扑空间关系。拓扑空间关系(topological spatial relationship,简写 topology)在地理信息系统及其空间数据模型的研究和应用中具有十分重要的意义。为此,下面专门用一个小节阐述之。

3　拓扑空间关系

Topology一词来自于希腊文,它的原意是"形状的研究"。拓扑特性指在拓扑变换(任意伸缩或变形,但不扭结或折叠)下能够保持不变的几何属性。地理空间关系的研究所特别关注的几个重要拓扑特性是连接性、包含和邻接性。

● 连接性。空间曲线的交点常称为结点。连接性(connectivity)指曲线或弧段在结点处的相互联接关系。在图2-3中,图a中的图形较大幅度地变形为图b所示,但是曲线或弧段的连接关系并不变化,例如,弧段1、2、3仍在结点*A*处连接;弧段2、5、6仍在结点*B*处连接。可见,"连接性"是一种拓扑特性。

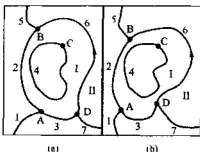


图2-3　拓扑关系示意

● 包含。包含(inclusion)关系有时仅就多边形或区域而言,但我们这里采用更广泛的理解,将点、线、面之间能够定义的所有包含关系或组成关系,都纳入包含的范畴。例如,一个点在一条线上或是其端点;一个点在一个区域或多边形的边界上或内部;一条线在一个多边形内部;一个区域在另一个区域的内部等。比较图2-3之a和b图,可以证实这些空间关系在拓扑变换中皆保持不变。

位于一个较大区域内部的小区域,常称为"岛"。例如,图2-3中,曲线4所围的小区域,就是曲线2、3、6所联合包围的大区域中的一个"岛"。

GIS中有两种基本组成关系值得注意。其一,一条线可以看作由一连串点构成。如上述,地理信息系统中的曲线,实际上是由很多足够小的线段组成的。因此,计算机中只要给出一条曲线所有线段之结点的坐标系列,就可以确定该曲线。其二,一个区域或一个多边形由若干条线围成,这种情况也称为"多边形区域定义"。例如,在图2-3中,曲线2、3、4和6围成了一个环状区域,它们定义了多边形Ⅰ。在拓扑变换中,尽管形状变化,但这个环状区域由曲线2、3、4和6围成这一特性也是不变的。

● 邻接性。邻接性(contiguity)指共有公共边的两个区域的邻接关系,如图2-3中,以曲线6为公共边的多边形Ⅰ和Ⅱ,就是相邻接的区域。这种邻接性也不会因拓扑变换而改变。

如果给公共边指定方向,邻接性还可以进一步描述为"右邻"还是"左邻"。例如,图2-3中,假定我们指定曲线6采取图中所示的方向,那么,顺该方向看,多边形Ⅰ就是左邻,或左多边形,多边形Ⅱ就是右邻,或右多边形。如果曲线6的走向反转,那么,左邻与右邻关系也要颠倒过来。

2.1.4 有关地理空间数据结构和模型学习的几点总体说明

上面讲述了地理空间数据组织的主要对象，也就是从“概念模型”的层次上讨论了GIS从地理空间世界到计算机世界的思路。下面应当向“地理空间数据模型”和“数据结构”的层次深入，具体地讨论地理空间数据怎样在计算机里组织和表达。下一节（本章第2节）将开始正式展开那些内容。但是，在此之前，我们觉得还有几个问题应交代一下，这些问题从总体上关系到本章2～4节的学习。

1. 关于空间数据结构和空间数据模型两个概念之间的关系

首先说明，从现在开始，为叙述简便，我们将如同众多GIS学者一样，在“地理空间数据结构（或模型或建模）”用词频繁的场合，将其中的“地理”两字省略掉，即简化地称为空间数据结构、空间数据模型和空间数据。

空间数据结构和空间数据模型研究地理空间数据组织和管理。二者之间的关系，与一般的数据结构和数据模型的关系有两点相似之处。其一，空间数据结构所做的数据组织工作，比空间数据模型更“基层”一些，它偏重数据表达的物理实现，而空间数据模型涉及到空间数据管理的层次。其二，同普通数据的数据模型一样，空间数据模型的命名通常也与相应的空间数据结构相同，例如矢量数据模型对应着矢量数据结构；栅格数据模型基于栅格数据结构等。

但是，空间数据结构和空间数据模型两个概念之间的关系，与普通数据的数据结构和数据模型又不完全等同。总的说来，它们之间的界限更不清楚，重叠的程度更大些（有兴趣探究其中差异者，可参阅陈述彭等编著的《地理信息系统导论》）。究其原因，我们觉得还是来自于空间数据的特殊复杂性。地理空间数据既要表现地物的空间位置及其相互关系，又要保证一一对应地挂联地物的属性数据。因此，在构建空间数据结构时，必须考虑到如何模拟和表达地理空间事物和现象，因而必然涉及空间数据模型；反过来，在构架空间数据模型时，也必须考虑到具体的操作层次。

由此缘故，本书将不严格区分空间数据结构和空间数据模型两个概念，也不分专门讲述空间数据结构和空间数据模型。又考虑到本课程的性质，本书讲解空间数据模型和数据结构，着重于基本概念而不追求技术操作，所讲的空间数据结构，主要指空间数据的逻辑结构，一般不涉及逻辑结构以下的“数据存储结构”和“算法操作”层次。

2. 关于属性数据的组织

第二点需要总体性说明的是属性数据的数据组织问题。地理空间数据包括大量的属性数据。属性数据可分为两类：一类是说明空间对象本身性质的数据，大部分属性数据属于此类。这一类属性数据的组织有一个基本要求，就是

要保证它们能与相应的空间对象相挂联。另一类属性数据说明其他的附加信息，如分析统计结果以及为输入输出服务的数据等，它们并不一定属于某一个具体的空间对象。

属性数据可以采取数字和文字等形式。数字形式的属性数据，如污染物的排放量、道路的宽度等，在计算机中很容易组织；文字形式的属性数据，特别当其较长时，计算机处理就不甚方便。为此，文字形式的属性数据常常被转化为数字形式来处理。一种最常见的例子是建立各种分类系统的“编码系统”，即用一系列数字来代表文字类型。例如，在土地资源管理信息系统中，“土地利用现状”是一种基本的属性，每一地块的“土地利用现状”属性值，必须按照国家的土地利用现状分类系统来取定。在该分类系统中，国家不仅给出每一种土地类的名称和详细定义，还规定了每一种分类的标准编码，以方便于计算机运作。

表 2－1　土地利用现状分类(部分)

一级分类		2 级分类		3 级分类	
名　称	编　码	名　称	编　码	名　称	编　码
耕地	1	灌溉水田	11		
		望天田	12		
		水浇地	13		
		旱地	14		
		菜地	15		
园地	2	果园	21	柑橘园	211
				其他果园	212
		桑园	22		
		茶园	23		
		其他园地	24		
林地	3	有林地	31	用材林	311
				防护林	312
				经济林	313
				竹林	314
				薪炭林	315
				特殊用林	316
		灌木林	32		
		疏林地	33		
		未成林地	34		
		迹地	35		

一级分类		2 级分类	
名　称	编　码	名　称	编　码
牧场	4	(当地无)	
居民地及工矿用地	5	城镇	51
		农村居民点	52
		独立工矿	53
		特殊用地	55
交通用地	6	双轨铁路	61
		单轨铁路	62
		主要公路	63
		一般公路	64
		农村道路	65
		民用机场	66
水域	7	河流水面	71
		湖泊水面	72
		水库水面	73
		坑塘水面	74
		沟渠	77
		水工建筑	78
未利用地	8	荒草地	81
		裸土地	85
		裸岩	86

表 2－1 是我国江南某地区境内所包含的各种土地利用现状分类的简表，它只是国家土地利用分类中的一部分(地方土地管理局通常只关心自己境内所包含的类型)。表中可见，我国土地利用现状分为 8 大类(一级分类)；每一大类又分为若干二级类；一部分二级类还进一步分为三级类。一、二、三级土地利用

分类的编码分别为一位、两位和三位数；每一种下级分类的编码，为上一级编码加上一个数字。例如，林地为一个大类，编码为3；林地的二级分类的编码皆为"3"开头的两位数；而编码为31的"有林地"下的三级分类，编码皆为"31"再加上一个数字，如防护林编码为312等。显然，在建设土地资源信息系统及其数据库时，采用编码系统作为属性数据，要比使用文字名称方便得多。

很多领域、行业或学科都像土地管理部门一样，有着规整的，带有编码系统的标准分类系统。这些领域中的GIS应用系统，很自然地就按行业标准分类编码来组织属性数据。但也有些分类系统没有标准编码，或标准编码不十分适合于具体的应用，这时，GIS系统及其数据库的设计、开发者，就要设计出合适的编码系统来组织属性数据。

不过应当说明，采用所属领域的分类系统之编码作为计算机数据处理的方法，是一般信息系统和数据库系统等领域常用的方法，并非GIS独有的特点。

3. GIS可视化要求和传统美术的两种基本技法

第三个需要总体说明的问题，是"可视化"与构建空间数据结构与模型的关系。如§1.2.3所述，可视化是地理信息系统的基本要求，地理学及地图学是人类科学可视化的先驱领域，而地理信息系统从很大程度上说是地图学的延伸。不仅仅空间位置数据和空间关系必须可视化，属性数据也需要以符号或注记的方式在图面上表现出来（因图面负荷有限，通常只能表现部分属性数据）。地理信息系统从诞生之初，就把可视化功能作为必要条件和构思地理空间数据模型、结构的基本出发点。这是空间数据建模与普通数据建模的一个重要区别。

空间数据建模既然要求可视化，即利用图形图像手段来表现地理空间现象和事物，那么，它就必然与美术有相通之处。事实上，我们认为，美术家们表现周围世界的技法，对我们理解空间数据结构和模型的构建确实大有助益。

至今为止，美术主要是在二维平面上表现周围的三维世界。美术家的表现手法多种多样，但在表现的基本思路上，可归结为两种类型。其一可称为涂抹方式，油画、水粉画和摄影作品是其典型的例子。这种技法试图表现所观察到的世界的每个细部。因此，它涂抹画面上所有的地方，通过画面每个细部的明暗或色调来表现周围实体的界限和形态。例如油画中白色桌布背景上一个红苹果，桌布用某种白色油彩涂抹（非空白），白色油彩与表现苹果的不同层次红色调之间的界限，给出红苹果的外边界。另一种技法可称为线条方式，如卡通画、连环画、漫画和中国画中的工笔画。这种技法重在表现画家所关注的空间实体（包括其组成部分）的线条，而忽略线条以外的其他世界。因此，在这一类画上，除了所关注的实体及其组成部分的轮廓外，很多地方是完全空白的

线条画法的主要长处是简洁明快，便于表现主要矛盾和事物，但表现力不够雄厚，画面中较多非主要部分空白，以致主要事物缺乏烘托等。涂抹画表现

力较丰富,但较费工(摄影除外)。实践中,很多画家兼用二者,局部涂抹,局部线条,如中国水墨画、素描写生等。不论如何,涂抹法和线条法是人类在二维平面上表现周围立体世界的两种基本方式。

4 两种基本空间数据模型或结构,及其与计算机图形图像的关系

美术与空间数据模型二者虽然意向不同,一个以艺术美为目的,一个以科学真谛为目标,但它们都要可视化地表现周围世界,因而在基本表现技法上明显互通。空间数据模型经多年研究发展,虽已有较丰富内容,但其中两种最常用最基本的模型,即二维平面上的矢量数据模型和栅格数据模型,以及基于这两种模型的矢量型 GIS 技术和栅格型 GIS 技术,正分别与上述两种美术技法对应。在本章下面几节中将看到,矢量数据模型或矢量 GIS 技术对应于线条方式的美术技法,它只表现主要空间实体及其组成部分的轮廓线条;而栅格数据模型或栅格 GIS 技术对应于涂抹方式的美术技法,它将整个画面分为规整网格,并通过每个网格单元的明暗或色调来表现空间实体的界限和形态。

在计算机科学技术中,也有两个分支采用上述两种基本思路进行可视化表达。一种相当于线条方式的美术技法,一般称为计算机图形学(graphics)或技术;另一种相当于涂抹方式的美术技法,一般称为图像处理(image processing)。由于可视化表达的需要,地理信息系统与生俱来就与计算机图形图像处理技术结下不解之缘。GIS 在以自身的发展不断充实图形图像技术的同时,总是迅速吸取图形图像技术的新成果:矢量型 GIS 技术从图形技术吸取营养;栅格型 GIS 技术从图像技术吸取养分。由此缘故,在矢量型 GIS 领域,常应用"图形"这一名词。例如,将矢量 GIS 系统中的可视化表达成分称为"图形平台"。在栅格型 GIS 领域,则常采用"图像"这一名词,如遥感图像处理等。

但是应当指出,GIS 并不等同于图形图像技术,因为可视化并非 GIS 的全部,GIS 还要应用数据库等其他的计算机技术,来处理大量属性数据及其与空间位置特征的挂联问题,以进行空间分析,提取信息,获取知识等。因此,应当说,GIS"大"于图形图像技术,有很多图形图像技术所没有的内容。

下面,我们就来具体地讲解空间数据结构或模型。首先在第二节,讨论矢量数据结构或模型;然后在第三节讲授栅格数据结构或模型;第四节在二者基础上进行进一步提高的工作。

§2.2 矢量空间数据模型及结构

2.2.1 矢量空间数据模型

建立在二维平面上的矢量数据模型是目前 GIS 领域应用最广泛的、与传统

地图表达最为接近的空间数据模型。矢量模型采用相当于线条画的表达方式，即用点、线和多边形(闭合的线)来刻画所关注的空间对象的轮廓、空间位置及其几何关系，同时组织好属性数据，以便与空间特征数据共同描述地理事物及其相互联系。为了用计算机来实现这种表达，矢量数据模型通常包括下述构成成分(图2-4)：

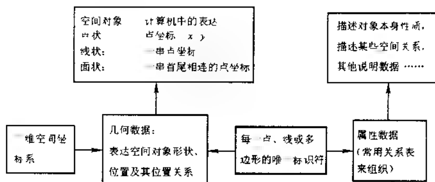


图2-4　矢量空间数据模型

1. 二维空间坐标系

建立在平面上的二维坐标系是表达空间对象的形状和空间位置的基础。GIS矢量数据模型中的二维空间坐标系可以是普通的平面坐标系，也可以是实际地表上的大地或地理坐标系。采用普通坐标系的地图也可以表达地物之间的相对布局，但归根结底还是要转换到地理坐标系中。地理坐标问题还将在本章下一节(栅格数据模型)中进一步阐释，并将在第四章第二节详细讨论。基于平面坐标系，矢量数据模型利用位置坐标(x,y)或者其组合来表达点状、线状和面状(多边形)空间对象的形状、位置和它们之间的空间几何关系。矢量数据模型中的坐标，原则上可任意取值，在精(密)度上也没有限制，这一点与必须等间隔取坐标值的栅格数据模型不同。

空间坐标(x,y)数据及其组合，称为几何数据。我们说过，地理空间数据是空间特征(或位置)数据与属性数据之和。这里的几何数据，就是空间特征数据的主要表现形式(少部分空间位置特征，如拓扑特征等也用属性数据来描述)。

空间坐标(x,y)是易于在计算机中运作的数据。

2. 几何数据

几何数据有时也称图形数据。在矢量数据模型中，模拟点状、线状和面状空间对象之形状、位置及其位置关系的几何数据分别组织如下：

• 对于本身的大小在研究中可以忽略的点状空间对象，用一个几何点坐标(x,y)来表达之。

- 对于本身宽度在研究中可以忽略的线状空间对象，用一连串点的坐标

$$(x_1,y_1),(x_2,y_2),(x_3,y_3),\cdots,(x_n,y_n)$$

来表达之。线状空间对象在荧屏、打印机、绘图仪等输出设备上输出时，输出设备会按计算机指令自动用线段连接相邻各点，将贮存在计算机中的这一串点坐标，可视化地表现一条线。不难理解，只要相邻点的密度足够高，就可以满意地模拟任意形状的曲线。

- 对于面状空间对象（区域），用首尾相连的一连串点坐标

$$(x_1,y_1),(x_2,y_2),(x_3,y_3),\cdots,(x_n,y_n),(x_1,y_1)$$

来表现其边界轮廓。可见，面状对象同线状对象一样，也是用一串坐标来描述，输出时也可视化地表现为一条线；但不同的是，这一串坐标的首尾是同一点坐标(x_1,y_1)，输出时表现为一串线段首尾相连，即闭合的线。而闭合的线，或首尾相连的一串线段，就是多边形。

3. 属性数据

属性数据通常用关系表的形式来组织。

属性数据可分为“基本属性数据”和“说明数据”两种类别。大多数属性数据都是基本属性数据。基本属性数据中，多数又是描述空间对象本身各种性质的数据，还有一些是描述某些空间位置关系和进一层空间关系的数据。注意，空间位置关系虽然主要靠几何数据及其可视化来表达，但有时也需要用属性数据来描述，例如一些拓扑特征等。进一层空间关系指基于位置关系的，空间对象间的其他内在联系，例如§1.4.1.4中的空间结构特征等。

说明数据是除基本属性数据外，为GIS组织或运作服务的属性数据。例如，一种典型的说明数据是描述空间对象的输出符号和注记的数据。正如传统地图一样，矢量数据模型的GIS在输出时，要用符号和注记来可视化地表现空间对象的主要属性特征。例如，点状对象在计算机中仅存储为一个几何点(x,y)及其链接的属性，但是在输出时，必须在(x,y)处用某种符号来表现该点状对象，如属性为“乡镇”的点用一个小圆圈来表示，“省城”的点可以绘两个相套的圆圈等。同样，线状对象可根据不同属性值，用不同颜色、粗细的实线或虚线来描绘，如干道用较粗的红线，乡村小道用细红线等。区域或面状对象则可以填充不同的图案和色彩。此外，点、线、面空间对象还需要文字注记，如河流线旁标注“黄河”等。显然，为了矢量模型的GIS系统能在具体点、线、面上赋予正确的符号和注记，必须给出相应的说明数据。

4. 唯一标识符

在矢量数据模型中，几何数据和属性数据常分别储存。为了保证§1.2.2.3中式(1.3)所述的空间特征数据与属性数据之间，落实于每个地物的一一对应的挂联关系，必须对所有点、线和多边形赋予唯一的标识符或标识码（identifier，常简

称为 ID)。一种常用的方法是,为每一点、线和多边形编号(不重复的唯一序号),并用之作为唯一标识符,例如,将某条线编为第 9 条线,"9"就是这条线及其属性的标识码;将某条多边形编为第 12 号多边形,"12"就是这个多边形及其属性的标识码等。基于唯一标识符的链接在计算机中可以通过指针等方式来实现。

以上四种成分构成了 GIS 的矢量空间数据模型的一般内容。应当指出,这些内容中有部分与一般计算机图形处理技术相通;但是,地理属性数据及其通过唯一标识符与几何数据的挂联关系,以及地理坐标转换,是一般计算机图形处理所没有的 GIS 的特色内容。今后我们还将讲到矢量 GIS 与计算机图形学的其他区别。

上面介绍了矢量数据模型的内容。在这个总框架下,还有不同的矢量数据组织方式,或不同的数据结构。下面将给出两种有代表性的数据结构,一个较简便,一个较完备。讲解这两种结构不仅有实用意义,而且有利于理解基本概念。

2.2.2 有代表性的矢量空间数据结构

1. Spaghetti 结构

Spaghetti 结构是较简便的矢量数据结构,早期的 GIS 软件,以及现在的一些桌面绘图或制图系统常采用这种结构。Spaghetti 结构主要面向多边形来组织数据,并将多边形边界看做是线的简单闭合;不从属于任何多边形的线和点才另外组织。因此,这种方法有时称为环状多边形数据结构。

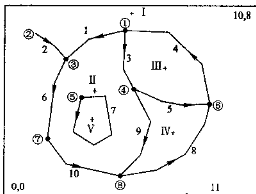


图 2-5 矢量数据结构示意

参见图 2-5,该图不仅用于本小节,也作为下一节讲述拓扑数据结构的示例,以便读者比较这两种结构。为此,有关该图的一些背景和定义应说明一下。首先,该图的坐标系原点位于图框的左下角点,图框的右上角点的坐标为 10,8 由此,图中每一点的坐标不难大体估计。其次,为便于说明问题,图中每条线由

不多的点构成。其中,线的端点(不少端点即线之交点)称为结点(node),其他的点称为中间点、节点或顶点(vertex)。每段线的序号用不带圈的阿拉伯数字表示;带圈的数字则表示结点;多边形的序号由罗马数字表示。

在 Spaghetti 结构中,除了不属于任何多边形的线和点外,地理空间数据基于多边形来组织,组织顺序通常为:多边形序号、体现多边形边界的一串点坐标,及该多边形的属性数据。例如,3(Ⅲ)号多边形的数据可以组织为:

3(多边形序号);4.7,7.0;4.7,5.3;5.1,4.5;6.2,3.9;7.1,3.8;8.1,3.9;7.8,5.5;6.5,6.8;4.7,7.0(多边形边界的一串点坐标);属性数据。

上述坐标串即图2-5中3、4和5号线的所有点的坐标串,其首尾坐标都是(4.7,7.0)。同理可知,4(Ⅳ)号多边形数据中的坐标串应是图中5、8和9三条线所有点的首尾相连的坐标串;2(Ⅱ)号多边形的数据则要包括1、3、6、9和10号五条线的所有点的坐标串,如此等等。不属于任何多边形的点和线(如图2-5中的2号线),另外组织,一般也采取点或线序号、坐标或坐标串,及属性数据的组织顺序。

Spaghetti 结构的优点是简便,易于实现以多边形为单位的运作,数据输入后无须进行大量编辑整理,即可方便地显示,能适应某些桌面绘图或制图系统的需要。但它也有若干不足。第一,每一条多边形的公共边界线要记录两遍例如图2-5中,记录2、3号多边形时,3号线会记录两次,记录2、4号多边形时,9号线会记录两次。这将导致较大的数据冗余,及制图时重复绘制的问题。第二,不能表达边界和多边形之间,及相邻多边形之间的拓扑空间关系。这是 Spaghetti 结构的两点主要不足。

另外的不足是,其一,难以表达多边形包含关系,例如图2-5中的2和5号两多边形有包含关系,但在 Spaghetti 结构中都作为多边形一样看待、编码,以致在 Spaghetti 结构中很难判断2号包含5号多边形。其二,采用 Spaghetti 结构的软件的数据输入方法通常是每个多边形周边输入一圈,这样,多边形公共边界线的两次输入记录常不一致,造成人为的间隙。这两点不足可以通过软件上努力来弥补。

2. 拓扑空间数据结构

拓扑空间数据结构是目前较完备的矢量数据结构,其数据组织的要点如下:

- 线(line 或 arc)由结点(node)定义,结点或者是一条线的尽端,如图2-5中的结点②;或者是线的交点,包括一条线自身闭合时的交点,如图中结点⑤。这就与 Spaghetti 结构有明显不同。在 Spaghetti 结构中,多边形的周边被看做是一条闭合线,例如包围3号多边形的闭合线在该结构中实际上没有3、4和5号线之分(上面用3、4和5号线来描述3号多边形的边界,是为了帮助理解 Spa-

ghetti 结构重复记录或编码的问题)；而在拓扑结构中，3、4 和 5 号是被结点分开的三条线。同理，6、10 号线是被结点⑦分开的两条线。

● 每个多边形由一个外环，以及零个或多个内环组成，详见下节阐释。

● 线是有方向的，在线的数据中，按线的方向，列出包含组成它的结点和所有中间点的坐标序列。

● 多边形的数据只记录围成它的若干条线的线号，而不再列出其边界坐标串；但是，每个多边形内部设置一个唯一的标签(label，如图 2-5 中的各“+”号表示的点)，标签点的坐标含在多边形的数据中；多边形的属性数据将挂联到该标签上。

● 独立点状地物的数据(点坐标和其属性数据)单独组织。

● 点、线和多边形之间的拓扑关系用若干属性数据来描述。其方法如下一小节所述。

3 拓扑关系及其表达

拓扑矢量数据结构中表达的拓扑关系主要有三种。

● 线拓扑。线拓扑体现线与其结点的联结关系，和以其为公共边的两个多边形的邻接关系。数据组织的一般形式是：在一条线的数据中，列出该条线的“起始结点”和“终结点”的序号，以表现该线与其结点的联结关系，如图 2-5 中 9 号线的起始结点为④，终结点为⑧；另方面，列出该条线左面的多边形和右多边形的序号，以表现该两个多边形在这条线两侧的邻接关系，这里，左和右是从该线方向看去的左和右，如图中 3 号线的左多边形和右多边形分别是Ⅲ和Ⅱ号多边形。

由此可体会到拓扑矢量数据结构中“线的方向”之意义：线没有方向，就谈不上起始结点和终结点，也不能确定左多边形和右多边形；线没有方向，更谈不上是“矢量”。因此，从严格意义上说，只有像拓扑结构这样的数据结构才是“矢量”数据结构。

● 结点拓扑。表现在该结点上的各线的联结关系。数据组织的常用形式是，在一个点的数据中，给出交于该点的各条线的线号，例如图中交于结点①的，是 1、3 和 4 号线；同时，在线序号前加上正负号，以表示该条线是起始于此点，还是终止于此点。

● 多边形拓扑。表现多边形与围成其边界的诸线条的构成关系，也能表现多边形之间的包含关系，即“岛”关系(如果有的话)。数据组织的常用形式是，在一个多边形数据中，列出构成其边界的各条线的序号。例如图 2-5 中，Ⅴ号多边形是Ⅱ号多边形所包含的“岛”；围成Ⅴ号多边形的是 7 号线，围成Ⅱ号多边形的各条线的序号是 1、3、6、9 和 10(外环)，以及 7 号(内环)。内、外环关系还可用某些专门编码来标识。此外，多边形拓扑常在线序号前加上正负号，以

表示该条线围绕该多边形是顺时针还是反时针方向。

4. 讨论

综上所述,拓扑空间数据结构确实是较完备的矢量数据结构。一方面。它较好地体现了地理空间点对象之间、线对象之间、多边形对象之间,以及点、线和多边形对象相互之间的几何关系,从而可能进行基于这些关系的多种空间操作或分析,例如邻域分析、空间选择、剪切和叠置分析等。另一方面,拓扑空间数据结构中,点、线、面之间构架清晰,没有重叠和漏洞,主要的坐标点数据皆组织在线数据中,而多边形由若干线条装配而成,因而不再需要包含坐标点数据(除标签点外),整个数据组织得简练而干净,基本上消除了冗余。

但是,拓扑数据结构也有缺点。正是因为结构完备,进行某些简单的操作(如显示)反而不简便,特别是拓扑结构对数据有一系列严格要求,因而数据准备很费功夫。如图 2-6 所示,其中图 a 表示城市道路网的一个局部,图 b 是图 a 中的虚线框部分的放大形象(计算机中放大缩小很容易)。在图 a 中可见,街区Ⅰ至Ⅴ被 1、2 和 3 号道路分开,看上去都是完整多边形;但是一旦放大,即在(b)中,就发现 1、3 号线并未真正咬合,以致街区Ⅰ和Ⅱ在计算机中实际上是一个多边形。这样,街区Ⅰ和Ⅱ既不可能有正确的拓扑关系,也不能挂联各自的属性数据。同样,看上去相交的 2、3 号线对计算机来说,只是两串(x,y)的坐标,在计算出这两条线的交点坐标以前,计算机并不知道这两条线是相交的,因而也就无法从性质上区分Ⅰ、Ⅲ、Ⅳ和Ⅴ几个街区,更谈不上表达拓扑关系。因此,这样的数据不符合拓扑数据结构的要求。

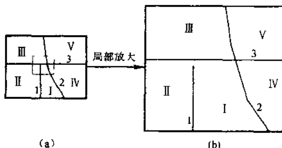


图 2-6　看似闭合的多边形和相交的线

但是,图 2-6 的情况对于只要求能较好显示图形的 Spaghetti 结构来说,则是允许的。由此可见,同 Spaghetti 结构相比,拓扑数据结构在数据准备过程中要做多得多的编辑工作,包括计算各条线的交点和真正封闭多边形等。拓扑数据结构所要进行的数据编辑工作,将在下一小节中进一步说明。由于拓扑数据结构在数据准备上的繁难性,在一些空间分析要求不高的场合,人们仍然愿意采用 Spaghetti 结构等较简单的数据结构。

除拓扑数据结构和 Spaghetti 结构外，还有别的矢量数据结构，本书不拟介绍，因为其他矢量数据结构一般可以看做是拓扑数据结构和 Spaghetti 结构的中间类型，且现在已较少应用。

2.2.3 不同格式的拓扑结构数据实例

1. 不同格式的拓扑结构数据

前面讲过，在采用同样空间数据结构的情况下，不同 GIS 软件商或数据供应商还可能采用各自的数据格式或文件格式。国际上采用拓扑数据结构比较有名的软件有美国 ESRI 公司（Environmental System Research Institute, Inc.）的 Arc/Info 软件和 Intergpagh 公司的 MGF（Module GIS Environment）软件等；比较有名的拓扑数据供应商有美国人口调查局（Census Bureau）和美国地质测量所 USGS（U. S. Geological Survey，有时也译为调查局）。美国人口调查局建立的双重独立地图编码系统 DIME（Dual Independent Map Encoding）是世界上最早的拓扑数据，其提供的人口调查数据（TIGER）成为很多社会科学研究的基础。大多数 GIS 书籍都有关于 DIME 和 TIGER 的介绍，读者很容易找到并参阅之。

USGS 是美国地形图的主要提供者，它所提供的数字产品中有一种常称为 DLG（digital line graphics，译为数字线划图）格式的数据。该数据特别适宜作为拓扑数据结构的教学实例，同时还有利于帮助学生理解“数据格式”概念及其与“数据结构”概念的关系。下面就以它为例讨论之。

2. DLG 数据格式及其实例

DLG 数据以文本文件的形式提供。仍以图 2-5 为例，该图的 DLG 数据如表 2-2 所示。由于该图尚未赋属性，数据中有很多值为 0。我们仅需专注于这里 DLG 数据的非零部分，以便具体理解 DLG 数据的格式及其拓扑表达。

DLG 数据依次由四个部分组成：

- 第一部分。“文件头”，1 至 15 行，描述该 DLG 数据的生产日期、地点、人员、经纬度范围等。

- 第二部分。从第 16 行起为各结点（node）和独立地物点的数据，每一点的数据占 2 行，皆以字母 N 开头，其非零数据依次为：点序号、点的坐标、以该点为端点的线的数目和序号（若为终结点，该线号为负）。例如，上述资料表明（参照图 2-5），这里共有 8 个结点，其中第 1 号结点（以 N 和 1 开头的数据）的坐标为（4.7，7.0），有 3 条线交于该点，它们分别是 4、3 和 1 号线，其中，3 和 1 号线出发于该点，4 号线终止于该点（线号为 -4）。

- 第三部分。紧接结点数据，为多边形或区域（area）数据。每一多边形的数据占 2 行，皆以字母 A 开头，其非零数据依次为：多边形序号、其标签点的坐标、围成该多边形的线的数目和各条线的序号，其中反时针转的线的序号为负。

例如,第4号多边形(以A和4开头的数据)的标签点坐标为(6.7,2 6),该多边形由3条线围成,它们分别是5、8和9号线,其中,5号顺时针,8和9号线反时针转(线号为负)。

表 2-2 图 2-5 的 DLG 数据

(前15行文件头略)

```
N   ,     47     70     3     0   0
 4    3    ,
N   2     11     68     .     0   0
2
N   3     23     58     3     0   0
6    2    1
N   4     51     45     3     0   0
9    5    3
N   5     29     41     2     0   0
 7    7
N   6     81     39     3     0   0
 8    5    4
N   7     16     24     2     0   0
10    6
N   8     45     08     3     0   0
-10    9    8
A   1     50     76   0   7   0   0   0   0
-11   0   1   6   10   8   4
A   2     35     44   0   7   0   0   0   0
3    9   10   6   1   0   7
A   3     63     54   0   3   0   0   0   0
 4    5    3
A   4     67     26   0   3   0   0   0   0
5    8    9
A   5     34     32   0   1   0   0   0   0
 7
L   1   1   3   2   1       3   0   0
47      70      33      66      23      58
L   2   2   3   1   1       3   0   0
11      68      19      63      23      58
L   3   1   4   3   2       3   0   0
47      70      47      53      51      45
L   4   6   1   3   1       4   0   0
81      39      78      55      65      68
47      7.0
L   5   4   6   3   4       4   0   0
51      45      62      39      71      38
81      39
L   6   3   7   2   1       3   0   0
23      58      19      49      16      24
L   7   5   5   5   2       6   0   0
29      41      26      27      34      22
42      26      39      42      29      41
L   8   8   6   4   1       5   0   0
45      08      60      10      71      17
79      29      81      39
L   9   4   8   4   2       4   0   0
51      45      56      31      53      17
45      08
L  10   7   8   2   1       4   0   0
16      24      23      14      45      08
L  11  (略)
```

但是，若多边形内有"岛"，则用"0"将多边形的外环和内环边界上的线的序号分开，相应地，围成该多边形的线的数目加1；且内环方向反算。例如上例中，可以发现1号和2号多边形内的边界线序号中皆有"0"，说明其内部含有岛。其中，1号多边形的外边界或外环为11号线（反时针，负号），内边界或内环由1、6、10、8和4号线（反时针取正号）构成。2号多边形的外边界由3、9、10、6和1号线（后三者反时针），内边界为7号线（反时针取正号）。

● 第四部分：线（line）数据，皆以字母L开头。这是DLG中数据量最多的部分，因为这里包含众多的点坐标（图2-5中特意少绘制各条线的中间点以便于讲解）。每条线的非零数据依次为：线序号、起始点序号、终结点序号，左多边形序号、右多边形序号、所含的点（包括两端点）的数目和坐标序列。例如上例中，第6号线（L和6开头的数据）从3号点起始，在7号点终结，左边为2号多边形，右边为1号多边形，含3个点，其坐标为（2.3,5.8），（1.9,4.9）和（1.6,2.4）等。

3. Arc/Info的拓扑数据表达

Arc/Info多年来一直是主导GIS技术潮流的最重要的大型GIS软件，其数据模型和结构包含较丰富的内容，这里仅介绍其中有关拓扑结构的部分，着重比较它与DLG数据格式不同的特点。在DLG数据中，每一点、线、面的空间位置数据（坐标和拓扑数据）与属性数据放在一起。如果我们给图2-5中的线和多边形赋以（自然或人文）属性，例如线为编码等于62或63的道路，每条路宽度为10或15m等，那么，上述DLG中的一些为"0"数据就会为62或63、10或15等所代替。位置数据与属性数据放在一起的这种特点在其他数据供应商的数据也时常可见。

但是，GIS软件商的数据通常不能这样组织。在Arc/Info中，几何坐标数据采取多种存储形式，而属性数据和部分拓扑数据放在关系表中另行存储。仍以图2-5中的图形为例，假定图中部分线条为编码63、路宽为10m的道路，另一部分为编码62、路宽为15m的道路，那么，图2-5中各条线在Arc/Info中的属性表将如表2-3所示。为方便起见，表2-3中的字段相对Arc/Info原表的字段略有简化。这里给出的各字段的意义依次为：线的序号、线的起始点、终结点，左多边形、右多边形、线的长度和若干地理属性（这里为道路类型和路宽）。读者可参照图2-5和上述DLG数据检验各条线的拓扑关系。正因为拓扑数据与地理属性数据放在一个关系表中，人们有时也把拓扑数据看做是属性数据（称为拓扑属性），尽管拓扑数据实际上是空间位置数据的一部分。

在Arc/Info中，也有可供用户打开和应用的多边形属性表和点属性表（独立点或点状地物的属性表）。点、线和多边形的属性表中皆有专门的序号（称为内部ID号，多数情况下与表头的记录号一致）作为每一点、线和多边形的唯一

标识符。通过它,每一点、线和多边形的属性一一对应地链接到其坐标数据上。值得一提的是,Arc/Info 仅在线属性表中给出拓扑属性供用户应用;点拓扑和多边形拓扑在供用户应用的点属性表和多边形属性表中反映不出。

表2-3 图2-5各线在Arc/Info中的属性表

No	Fnode	Tnode	Lpoly	Rpoly	Length	Type	路宽
1	1	3	2	1	2.74	63	10
2	2	3	1	1	1.59	63	10
3	1	4	3	2	2.58	63	10
4	6	1	3	1	5.27	62	15
5	4	6	3	4	3.19	62	15
6	3	7	2	1	3.48	62	15
7	5	5	5	2	5.89	62	15
8	8	6	4	1	5.26	63	10
9	4	8	4	2	4.20	63	10
10	7	8	2	1	3.55	62	15

4. 再谈数据结构和文件(或数据)格式

上述 DLG 和 Arc/Info 实例及其讨论清楚地表明,同样是采取拓扑数据结构,不同数据商和软件商还会有自己具体的数据组织形式。这就是文件(或数据)格式。事实上,上述实例不仅是为了帮助认识拓扑数据结构,而且也是为了对"数据结构"和"文件(或数据)格式"这两层概念有一个具体的认识。

这一认识具有重要的意义。在地理信息系统实践中,经常遇到需要将不同数据结构或格式的软件平台集成应用的情况。这个问题也即现在整个计算机领域都甚为关注的"互操作性"(interoperability)问题。可惜的是,由于地理空间数据的特殊复杂性,同一般计算机应用领域相比,GIS 领域在互操作性方面的进展要滞后甚多,GIS 实践中不同软件平台兼容并用的困难,不同数据结构或格式的数据相互转换的麻烦,一直是 GIS 领域发展的一个严重瓶颈,成为当前地理信息科学领域需要重点研究解决的课题之一。

2.2.4 矢量结构的数据的输入、编辑和输出

在了解了矢量空间数据结构的基本知识后,接下来自然想解决的问题是:这种数据结构在实践中怎样运作和实现?现实中的资料怎样变为计算机数据?又怎样从计算机输出,表达成人们习惯的形式?本小节就来解决这个问题

1. 矢量结构的数据的获取和输入

矢量结构的地理空间数据输入,包括空间位置数据和属性数据的输入两

者。文字、数字形式的属性数据之输入与一般计算机数据一样，不用解释。在空间位置数据中，拓扑等空间关系数据，通常是在数据输入以后进行数据编辑整理的结果。因此，这里需要着重谈的是空间坐标数据的输入。在这一方面，矢量数据结构一般采取下述四种途径：

- 人机交互，直接从计算机键盘或鼠标输入。这是一种最基本的输入形式，常常作为其他输入方式的辅助手段。例如，当需要补充一些点的(x,y)坐标时，可直接从键盘敲入；当需要在荧屏上添加、修改几何图形时，常用鼠标点绘；通过键盘或鼠标还可以指令软件平台自动给出所需图形，如椭圆或圆弧等。
- 直接采用现成的矢量结构电子数据。现成的矢量数据通常有两类。一类是软件商和数据商提供的数据。较流行的商用数据常能应用于多种软件平台，例如 Arc/Info 软件就能将 USGS 提供的 DLG 等商业数据转为自己的格式，直接应用。另一类是 GIS 系统的建设者自己利用现代化数字测量仪器采集到的地物坐标的电子数据，这些数据连同相关属性数据，作为数据文件一般能直接输入到 GIS 系统中。
- 从栅格数据文件转换而来。这一点将在本章后面讲述。
- 通过手扶跟踪数字化仪输入。这一手段现在虽然已用得不多，但它对理解矢量数据输入的概念特别有用，故这里稍用些篇幅。

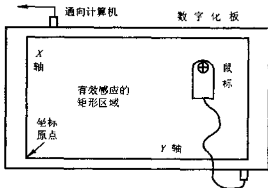


图 2－7 手扶跟踪数字化仪示意

手扶跟踪数字化仪（digitizer）是一种专门的矢量数据输入设备，常简称为数字化仪，其输入对象是地图等已有图形。数字化仪的主要构件是一块数字化板和一个与之用电线相连的鼠标（或游标，参见图 2－7）。数字化板内含有感应电路，板中有一个矩形区域为有效感应区，它构成了一个平面坐标系的第一象限。鼠标上有若干输入按键和一个透明的十字丝，只要鼠标贴于数字化板面，按输入键就能将鼠标十字丝中心所在位置的(x,y)通过电耦合机制送入计算机。实际输入时，将地图贴于数字化板面，使鼠标十字丝中心沿地图中的线

状地物(或区域边界)移动,并陆续按键,就会有表现该线条的一串(x,y)输入到计算机中。输入点状地物则只需使鼠标十字丝中心对准该点状地物,按一下键即可。

应当注意,经手扶跟踪数字化仪输入的地图点、线、面坐标,是数字化板平面坐标系的坐标,该坐标仅表现地图要素间的相对位置。若数字化仪为进口设备,其默认坐标单位常为英寸。这样的坐标常常需要转换成地理坐标,以便实际应用。为了进行这种转换,需选取若干个已知地理坐标的特殊点(例如地形图的图廓点等)作为控制点,通过这些点,建立数字化板坐标和地理坐标之间的数学转换关系(详见第四章第一节)。

2、矢量结构数据的前处理和编辑

输入计算机的矢量数据一般都有很多不足,如有重叠、短缺、不规范和不美观,以及尚无基本几何数据(长度、面积等)、拓扑结构等,因而还不能直接应用,需要经过一番处理,再加上属性数据,才能满足进一步的数据处理和分析的需要。这种为用户准备好基本达到要求的数据处理,有时称为数据的前处理或预处理(pre-processing)。注意,预处理是计算机数据处理的一环,并不包括计算机输入前的数据准备工作。

在矢量数据的前处理工作中,有些工作是通过计算机程序自动完成的;但更多的数据整理工作需要进行人机交互,即用户利用 GIS 软件平台的某些功能,通过人工干预来整理数据。这种人机交互整理矢量数据的工作,通常称为矢量数据编辑,其中大量的工作是几何图形的编辑。实际的前处理工作常是人机交互整理和程序自动处理反复结合的过程。

下面具体介绍常见的矢量结构数据前处理和编辑工作的内容。

3、一般矢量数据的前处理和编辑

本小节给出拓扑和非拓扑结构的矢量数据都需要进行的前处理和编辑,拓扑数据额外需要进行的工作将在下一小节中讲解。一般矢量数据的前处理主要有三类工作。

第一是基本度量计算。基本度量计算指空间度量关系的计算,如线的长度、多边形的周长与面积、两点间距离与方向、点到线的距离等的计算。这些计算通常由计算机程序自动完成,并已成为一般图形处理软件的现成功能。很多 GIS 软件在线和多边形的属性表中自动地给出线的长度和多边形的面积等。

第二是一般的图形编辑工作。一般的图形编辑工作大多需要在图形软件平台上进行人工干预。例如:

- 点、线、面实体的增加、删除、移动、拷贝、平移、旋转等。一般图形软件都有现成功能,用户通过鼠标等可方便地操作。
- 在线或多边形边界中删除、增加或移动其中间点(顶点或节点)。此编

辑的目的是改变线或多边形的形状。如图2－5中，第4号和第8号线的长度和弧度差不多，但8号线有3个中间点，4号线只有2个中间点，8号线就显得比4号线圆滑一点；不难想象，若4号线再去掉一个中间点，它的折线形态将更明显；反之，若在8号线每一线段外侧近处增加一个中间点，会使它更接近圆弧形。一般图形软件都具有让用户利用鼠标等工具增、删和移动中间点的功能。

• 曲线光滑处理，通过计算机程序自动增加中间点，使线或多边形边界变圆滑。其思路是，通过数学计算，寻求能拟合线条局部各点坐标的曲线函数，如2次或3次多项式；该函数找到后，就可以为该条线增加任意间隔的中间点。

• 曲线化简，相当于曲线光滑处理的反过程，其目的是在保持一条曲线基本形状的前提下，删节曲线中过密的中间点，以减小数据冗余。在地图向小比例尺缩编需要进行制图概括（综合）时，也需要曲线化简。不少GIS软件有曲线化简功能，但还不够理想。

• 点、线、多边形的连接、咬合或合并。

• 窗口裁剪，利用一个矩形框或多边形去裁剪图形。矩形窗口常用于局部图形的显示、放大和编辑等。更复杂的图形裁剪操作涉及到空间操作分析，不属于前处理的范畴。

• 图幅接边，是将两幅相邻地图边界上不一致之处吻合起来的工作，不仅要求两边地物在几何上吻合，在属性上也要一致。前者称为几何接边，后者称为逻辑接边。几乎所有的GIS软件都有图幅接边的功能，使用户能方便地通过人机交互进行图幅接边。

• 撤销与恢复编辑。

第三种前处理工作是属性数据输入。由于矢量结构数据需要做很多前处理及编辑工作，GIS的属性数据的输入常在编辑工作告一段落之后进行；但通常不可能一蹴而就，属性数据输入和图形编辑常常是一个反复交替的过程。

4．拓扑结构数据的前处理和编辑

前面§2.2.2.4已初步讨论过，同非拓扑数据相比，拓扑数据要额外做不少的前处理和编辑工作。参照图2－6和图2－8，现将这些额外的工作梳理如下。

• 让未到位的线咬合，应当接上的结点吻合（snap）。前者如图2－6中1、3号线；后者如图2－8a中的两个结点，如不咬合，本来是封闭的多边形将不成立。

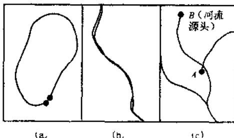


图2－8 需要进行数据清理的几个实例

● 计算线的交点，这些交点加上线条之端点构成拓扑数据中的结点，进而定义每一条线(或弧段)，以及结点和线之间的拓扑关系。例如，图2－6中的2和3号线的交点找到后，这两条线将被结点分割，重新定义。

● 用若干线或弧段装配并定义多边形，建立线和多边形之间的拓扑关系。

● 为每个多边形制造唯一的标签点，为此需要进行点在多边形内的判别。

● 建立点状、线状和多边形地物的属性表，表中给出某些基本度量计算结果(线长度、多边形面积等)或拓扑数据。

● 去掉多余的多边形。特别是重复数字化或不同来源地图叠合等原因，会导致同一地物有重叠弧线，如图2－8b所示。重叠弧线会形成一连串莫须有的细碎多边形，需要去除之。

● 清除不必要的秃结点和假结点。秃结点指弧线的端点，如图2－5中的结点②和图2－8c中的*B*点　自然界中有不少真实的地物表现为秃结点，如街道死胡同、河流源头等；但也有很多秃结点是多余的。例如，"丁"字型的接线在数字化输入时未接上，如图2－6b中的1号线，或过了头，如图2－8c中的*A*点，都会形成多余的秃结点。前者需要接上，后者需要剔除(一般而言，后者易于处理)。"假结点"则是两条线的交点，如图2　5中的结点⑦。假结点在多数情况下是多余的，但保留也无关大局，故并非所有的GIS软件都有处理假结点的功能。

上面例举的拓扑数据编辑内容中，多数可由计算机程序自动完成，拓扑关系可以自动建立。但是，即令如此，大量的人机交互工作仍不可免。例如，GIS软件可以剔除用户给定阈值的短线头和细碎多边形，但这并不能完全解决实际问题。这是因为，如果用户给定的阈值过大，可能剔除掉一些真实的短线头和细碎多边形；若阈值太小，又可能有部分该剔除的短线头和细碎多边形被保留下来。其他的拓扑编辑功能皆类似，都需要相当多的人工干预。

还有，上一小节谈到，矢量数据处理不可避免地会有大量图形编辑工作，而在拓扑型矢量数据的情况下，图形编辑还会引起更多的编辑工作量，因为每一点、线、面的修改都可能改变周边图形的拓扑关系，以致拓扑关系需要重新建立，某些属性需要增删或修改；修改后的数据还可能发生新的问题。由此缘故，准备完好、干净的拓扑数据的过程，往往需要经过若干次反复。

5. 矢量结构的数据的输出

GIS输出有两种主要形式：荧屏输出和硬拷贝输出。后者一般指通过绘图仪和打印机等设备在纸张上输出。现在真正以矢量数据形式输出的设备只有一种，即数字笔式绘图仪　它的绘图笔尖能够按程序指令自动地沿每一条线的坐标串，连接每两个相连的(x,y)，从而在图纸上画出线条。笔式绘图仪通常有八个不同颜色的绘图笔尖，能按地图设计的要求绘出黑白或彩色的图件　不

过，现在笔式绘图仪已用得愈来愈少，以栅格数据（下一节学习）形式输出的绘图设备，如喷墨绘图仪、喷墨打印机等，由于性能价格比迅速提高，已占绝对优势。荧屏输出也是采取栅格（点阵）形式。因此，目前在绝大多数情况下，矢量结构数据都是转换为栅格数据结构的形式输出的。本章下面两节将进一步解释这些问题。

在矢量数据输出方面，更值得注意的问题是地图符号和注记的输出。上面（§2.2.1.3）在讨论属性数据中的说明数据时，业已涉及过这个问题。事实上，如果没有符号和注记，矢量数据输出时只能是没有任何地理意义的点、线和多边形几何图形。为了将矢量数据输出成地图，必须用符号和注记来可视化地表现空间对象的主要属性特征。为此，矢量形式的 GIS 软件通常业已制作或准备好很多种表达点、线和面状地物的符号和注记素材，并有指挥栅格形式的输出设备动作的功能。符号包括点状符号、线状符号和面状符号。符号和注记一般也是矢量图形，如矩形、圆圈或其他图案，以及矢量汉字等。

在具体的矢量地图数据输出时，用户或者采用软件平台现有的符号和注记材料，或者再加工生成新符号；同时，还要给出 §2.2.1.3 提到的说明数据，具体指定哪一类地物（何种地理属性等）采用哪一种符号或注记，以及怎样实施等。这样，输出设备将在不同属性的点、线和面的位置上，分别绘出指定色调和图案的符号；并在指定的位置处绘出注记。例如，在某个居民点的（x,y）处绘一个小圆圈表示它是居民点；沿着某道路线的（x,y）坐标串，绘出红色虚线表示它是农村小路；在某河流旁的指定位置注记该河流是“黄河”等。

矢量数据的输出内容非常丰富，这里所写仅是为了帮助理解矢量数据的基本概念，详细内容将在第六章专门阐述。

§2.3 栅格空间数据模型及结构

2.3.1 栅格空间数据模型

建立在二维平面上的栅格（raster）数据模型采用相当于涂抹画的表达方式，或图像方式，通过画面上每个单元的明暗或色调来表现空间实体的界限和形态。为了用计算机来实现这种表达，GIS 栅格数据模型通常包括下述构成成分（参见图 2－9）：

1. 规整格网

首先，栅格数据模型将所研究的平面或地面区域划分为规整的格网，英文常称为 grid；格网的单元，即每一网格，通常为正方形，也可以是矩形、三角形、六边形等，英文常称为 pixel，有时也称 cell，中文一般称为像元（有时也称像素）。

格网划分为行和列，从上至下，分别为第一行，第二行，第三行，…，从左至右，分别为第一列，第二列，第三列，…等，如图 2－10 中 b 和 c 所示。

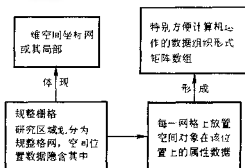


图2　9　栅格空间数据模型

在地学中，栅格数据模型的格网常常像地图一样，按左西右东、上北下南排列。由于卫星或其他飞行器上的遥感器（或传感器）不一定沿东西、南北方向对地面扫描，原始遥感图像的格网并不一定沿东西、南北向展开；但经过处理后，遥感图像常常是东西南北向的。

2　二维空间坐标系

建立在平面上的二维坐标系是表达地物形状和位置的基础。GIS 栅格数据模型主要采用两种直角坐标系。一种是一般计算机图像处理中常用的坐标系（参见图2－10b），其原点置于左上方，通常以左上角一行第一列的像元（中心或顶点）为原点，并以向右为 x 轴，向下为 y 轴，以“列距”Δx 和“行距”Δy 为坐标单位。这样，行、列数（整数）本身就是表达像元位置的坐标，例如第 i 列、第 j 行像元的坐标就是 (i,j)。这种方式来自于计算机图像处理从左到右、从上到下的习惯，与传统数学和传统地图的习惯不同。在地学处理中，当不涉及地理坐标，只需要体现研究地区内部单元之间的相对位置时，常采用这种图像坐标。

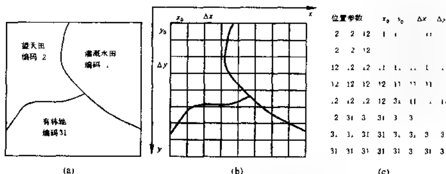


图 2－10　栅格空间数据模型示意

GIS 栅格数据模型常采用的另一种坐标，就是地理坐标。地学图像处理归根结底总是要落实到地理坐标上的，并采用与地理坐标轴平行的规整格网。注意，不同地理坐标系 x 轴和 y 轴的具体定义方法不尽相同。例如，地理经纬度

坐标习惯与数学相同，一般以向右（东）为 x 轴，向上（北）为 y 轴；而我国常用的高斯－克吕格大地坐标（相当于美国的 UTM 坐标），以向上（北）为 x 轴，向右（东）为 y 轴。

不同地理坐标系的具体定义虽然有差别，但有一基本点是共同的，即它们皆采取东西、南北向，而且在东西向，皆以向东（右）度量为正，在南北向，则皆以向北（上）度量为正。注意，这一点与上述计算机图像处理的坐标度量方向是不一致的。在 GIS 的栅格数据或图像处理中，常常既需要用到图像坐标，又需要转换为地理坐标。这就要求在应用栅格数据时，对不同坐标系之间的关系和转换要稍加小心。

当栅格数据模型之格网与地理坐标轴平行时，栅格数据的行、列数本身就体现地理坐标，这时，图像坐标与地理坐标的关系比较简单。参见图 2－10b，若网格或像元为矩形，那么，只要已知一个起始像元的坐标，例如第一行第一列的左上角像元中心的坐标（x_0，y_0），加上东西方向的"列距"Δx 和南北方向的"行距"Δy，任一像元（中心）的位置坐标就很容易得到。例如，第 i 列、第 j 行像元坐标为：

$$x_i = x_0 + (i-1) \qquad \Delta x, y_j = y_0 + (j-1)\Delta y$$

这里，Δy 在地理坐标系中取负值，因为行数增加，意味着向南（下），地理坐标数值减小。栅格数据模型中，多数情况下网格像元取正方形。这时，Δx 等于 Δy，计算更简便。

综上所述可见，栅格数据模型的规整格网本身就可以看做是一种直角坐标网，或坐标网的一部分；空间位置数据隐含在格网的行、列数之中，或者说，像元记录的顺序已经隐含了空间坐标；行、列数加上 4 个位置参数（图 2－10c），所有像元空间位置就可确定。但是，像元坐标取值是分立的或离散的，坐标值以 Δx 和 Δy 为单位等间隔地增减，而不能像矢量模型的坐标数据那样任意、连续地取值。这是栅格数据模型与矢量数据模型的一个明显差别。

3. 属性数据和分辨率

栅格数据模型中的属性数据是网格所代表的地面单元在某方面的特性，属性数据与空间位置数据的衔接是与生俱来的，因为在每一具体像元格子上放置的，就是相应的地面单元之属性数据。例如，图 2－10a 示意一个地区的土地利用现状，该地区含灌溉水田、望天田和有林地三种用地。现在我们用栅格数据模型来表达它。首先，我们采用我国土地利用现状分类的官方系统的编码作为属性数据（参见表 2－1），根据该编码系统，有林地、灌溉水田和望天田三种用地的分类代码，分别是 31、11 和 12。接下来我们将该地区划分为 8 × 8 的格网（图 2－10b），并给每一个网格或像元赋值，这个值就是网格（所对应的地面单元）的属性数据，即其所在的用地分区的地类编码。例如，位于格网下部的有林

地分区的各网格属性数据取值，应为有林地编码31。

有些网格位于分区边界上，我们取在该格子中占较大面积地类的代码，作为该网格的属性数据。例如，第六行第一列的网格，即像元(1,6)包含望天田和有林地两种用地类型，但望天田在像元中面积较大，我们就取望天田编码12作为该像元的属性数据。

由此可见，栅格数据模型中，像元是属性数据取值的最小空间单位；换言之，一个像元的某一种属性只能取一个值，该值实际上是像元所代表地表面积中占优势的那种属性值，或某种平均表达，像元内部的属性差异不能反映。由此也可见，网格或像元大小，决定了栅格数据模型中属性和空间坐标取值的精度。网格或像元所代表地面区域之线度的大小，称为栅格数据的"空间分辨率"，通常就简称为"分辨率"(resolution)。例如，若像元地面大小为30m×30m，分辨率就等于30m。栅格数据模型欲较精确地表达地理空间实体和空间对象，所取的格网必须足够地细密，分辨率就必须足够地高。

4. 矩阵数组

安置好每个网格的属性数据后，栅格数据模型中的数据就构成一种特别方便计算机运作的数据组织形式，即矩阵数组(图2－10c)。矩阵数组不仅方便计算机存储、显示或输出，而且可以用线性代数等数学方法来处理变换，分析计算结果还可形成新的矩阵数组。不论怎样变换，只要记住初始像元中心的坐标(x_0, y_0)和南北行距Δx和东西列距Δy，新矩阵数组每一个单元的空间位置始终在把握之中。

应当特别说明，上述栅格数据模型中的大部分内容，也适合于一般的计算机栅格图像数据；但地理坐标转换、地面分辨率和地理属性处理等内容，是GIS特色。栅格图像由于形象表现力较强，又方便于计算机处理，已被广泛地应用于多媒体、互联网和各种计算机输入输出设备等。栅格空间数据除了可以作为普通图像运作外，在GIS中的主要应用领域是遥感图像处理和地表三维立体表达。今后我们还将不止一次地谈到GIS栅格图像处理与一般的图像处理技术的区别。

上面讲述了栅格数据模型的一般内容。同矢量数据一样，在数据模型的总框架下还有不同的栅格数据结构。下面分别介绍栅格数据结构的两种主要类型。

2.3.2 完全栅格空间数据结构

完全栅格数据结构又称为简单栅格数据结构，它主要相对于压缩的栅格数据结构而言。完全栅格结构原封不动地采取规整的矩阵数组形式；而压缩结构则将矩阵用某种精简方式来记录(图2－10c中很多像元属性值相同，完全可以

采取简化的记录方法）压缩的栅格数据结构将在下一小节详细讨论，本小节讨论完全栅格结构的具体特点。

1. 数据量及其与精度的关系

由于原封不动地采取规整的矩阵数组形式，完全栅格数据结构的数据量容易由下式计算：

数据量（字节）= 像元总数 × 每个像元（属性数据）所占字节数

= 行数 × 列数 × 每个像元所占字节数

在不同具体格式的栅格数据中，每个像元属性数据所占字节数不尽相同。绝大多数格式的栅格数据采用整数属性数据，而且其中又有多数将属性值限定在 0 ~ 255 的整数范围内，即每一个像元占用一个字节。这样的栅格数据之数据量（字节）就等于像元总数。例如，一个分辨率为 10m 覆盖 1 万 km（100km × 100km）地区的栅格图像，应有 10 000 行 × 10 000 列，含 1 亿个像元，即数据量为 100M 字节。换言之，这样一个图像就要占用 100M 字节的存储空间。若分辨率再增高，或覆盖地区再增加，或图像是多波段的，数据量还将成倍增加，不仅可能占据过大的存储空间，而且会导致处理时间或计算效率上的损失。

因此，在采用栅格数据模型时，首先要处理的主要矛盾是数据量与精度的关系。为此，一要选择适合的空间分辨率：分辨率低，精度低，地面表达粗糙；分辨率过高，数据量会过大。二是尽可能设法压缩数据（§2.2.3）。

2. 不同地学应用类型的栅格数据

栅格图像有多种不同格式，常用的如 BMP、GIF、JPG、PCX、TIF 等。不过，本小节要讲的"类型"不是指数据格式，而是指地学上，从不同角度表达地理空间和实体的栅格数据。作者认为，这样讲解有利于学习理解，尽管一般 GIS 书籍上没有这样的编排。

第一类：表达线状和点状空间对象的栅格数据

这里，线状对象包括线状地物和多边形边界。在栅格数据中，点由一个非零值像元表示；线由一串彼此相连的非零值像元构成。点、线像元以外的像元值为零。点、线像元的值可取其地理属性值，如道路或河流的编码、等高线的高程值等，也可简单地取为"1"，如图 2-11 所示。像元数值仅取"0"和"1"的图像称为二值图像。

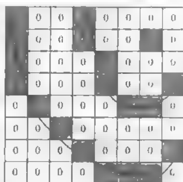


图 2-11 点状和线状空间对象的栅格数据

二值图像是最精简的栅格图像，它表达每个像元处"有"或"没有"某种属

性。但是,如果加上少许说明数据,二值图像也能表现更多的内容。例如,假定我们注明图 2-11 中,从第四列第一行像元(4,1)开始的线是高程为 300 的等高线,那么,我们完全可以利用二值图像,而不需要让这条线上每个像元都取值为 300,增加冗余。

图 2-11 的线条形象是栅格数据表达线状对象的标准形式。在不少情形下,例如从遥感图像中提取出的线状地物,或扫描仪扫出的地图线条常常有不止一个像元的宽度,就像多列的行军队伍。后者是因为,地图上线条本身宽度通常比扫描仪感应像点还大(数字扫描仪的原理将在本节后面介绍)。这种多列行军队伍般的宽线条,通常需要加以细化,形成图 2-11 所示的像元间单一连接的线条。"单一连接"的定义是:第一,线条不中断,线上每两个像元间必须连接,或邻边相连,或顶角相连;第二,任何三个像元不得两两间都相邻。

第二类:表达面状(区域)空间对象的栅格数据。

表达面状空间对象的栅格数据有时(特别在遥感研究中)称为分类图像,因为多边形或面状空间对象其实就是有某种共同属性的一个区域。分类图像通常采用所涉及的分类系统的编码作为属性数据值。图 2-10c 就是一幅较典型的表达面状空间对象的栅格图像。由于通常包含多种类别的分区,分类图像很少采取二值图像的形式。

第三类:实录型栅格数据。

上面谈的是表达点、线、面状地物的栅格图像,如 §2.1.3 所述,点、线、面状地物是人类对地理空间事物简化抽象的结果。还有一种栅格图像如同摄影照片那样,是周围世界本来面目的实录,相当于原始素材图像,尚未经抽象概括。这种图像主要有两种来源。一种是通过数字扫描仪将照片等硬拷贝图或图片,扫描成栅格数据。另一种是通过数字感应或采集设备,如各种遥感器(sensor)或数字摄像机等,直接采集周围景观的数字图像,或者将普通摄像机的图像转换成数字图像。现代在多媒体和互联网上运作的大量图像,都是实录型栅格图像。

对我们来说,最重要的实录型栅格图像是通过遥感器采集的地学遥感图像。地学遥感设备通常置于卫星或飞机等飞行器上,对地面扫描感应,均匀地、依次地记录下一个个地面单元在某个或某些电磁波段的"亮度"(来自地面单元的辐射强度值),形成栅格遥感图像。通过计算机从遥感图像中提取各种地物信息的学科,称为遥感图像处理。遥感图像处理中最常见的信息提取工作,就是从实录景观的遥感图像中,按特定的学科角度,区分并提取出同类型的面状地物、线状地物和点状地物。当遥感图像被分为若干分区时,例如土地类型分区或植被分区时,就形成像图 2-10a 和 c 那种形态的分类图像;当从遥感图像中提取出线状和点状地物时,可能形成如图 2-11 那种形态的栅格图像。

3. 多重属性下的栅格数据结构

接下来的问题是，栅格图像中一个格子只能放置一个数值，即一个栅格图像只能表达一种因子属性（如地类或植被），而实践中经常涉及多重因子属性的研究，进行多因子综合分析等。栅格数据结构解决这个问题的主要办法是：采用多层（layer）栅格数据，每一层栅格数据采用完全一致的格网，分别表达某一种因子属性。注意，矢量数据也常需要采用多层表达方式，例如对同一地区绘制各种专题地图。但应当看到，一幅矢量地图可以表达不止一种因子属性（例如交通、水系等属性可以同时绘于一幅图中），而栅格图像每一幅只能表达一种因子属性。因此，当涉及不止一种因子属性的分析研究时，栅格数据更加需要多层表达方式。

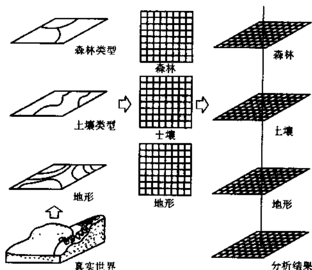


图 2-12 多层（多重属性）栅格数据示例

在 GIS 实践中经常遇到的多层栅格数据主要有两种类型。其一是多因子综合分析。在地理、资源、环境等很多领域所进行的区域分析，绝大部分涉及到多重因子。图 2-12 给出一个多因子分析的三层栅格数据示例，其中森林（植被）、土壤和地形是环境研究中常用的几个因子。图左部表示从真实世界的一个区域抽象出森林、土壤和地形三种属性的地图，再用完全一致的格网将这三重属性表达成栅格数据模型，从而可进行图右部所示的这三个因子的综合分析或叠合分析，最后用同样格网的一幅新栅格图像表达分析结果。

另一种常用的多层栅格数据是多波段遥感数据。遥感图像通常是多波段的，是若干不同波段的遥感设备同时对同一地区进行扫描感应的结果。例如，最常用的美国陆地卫星遥感数据，即分辨率为 30m 的 TM 数据，含 7 个波段的

栅格图像数据(其中6个波段的格网完全一致),是TM遥感器从红、绿、蓝,以及4个不同的红外波段同时对地面扫描的结果。由于各种地物在不同波段的电磁辐射和反射的特性各不相同,多波段遥感数据所包含的地物信息,比单一图像要多得多。大量的遥感图像处理工作,就是对多波段的,即多层遥感数据进行综合分析。

4. 多层栅格数据结构的文件格式

多层的完全栅格数据结构可以采取不同的文件格式。其中一种常用的格式是每一层栅格图像按正常行列序形成一个文件,一个接一个文件存放;这种方式在遥感图像处理中称为波段顺序(BSQ)格式。另一种常用格式是逐行格式(BIL)。在BIL数据中,所有波段逐行交叉排列,共同形成一个文件。该文件的第1行是第一波段第1行;接下第2行是第二波段的第1行;再接下是第三波段第1行,…… 所有波段的第1行依次排好后,再依次排列所有波段的第2行,然后依次排列所有波段的第3行,……如此等等。这种格式有利于多重图像的叠合分析或操作,因为同一行像元的多重属性值的物理地址相近。

多重属性的栅格数据也可用另一种格式来组织,即采取一个栅格图像加上属性表的方式。该属性表中的每一行记录着每一像元的多重属性;而在该栅格图像中,像元的值则是指向属性表中相应记录的指针。这样,也比较容易找到任何像元的多重属性值。

2.3.3 压缩的栅格空间数据结构

压缩栅格数据的必要性从图2-10和图2-11所示的栅格数据一目了然,在那里,很多像元取相同数值,数据冗余明显。压缩的栅格数据结构就是对完全栅格数据结构采取简化方法来记录的结果。下面介绍几种不同的压缩方法导致的压缩结构。

1. 游程长度编码或块码

游程长度编码和块码适宜于压缩像元属性值连接的、成片的相同的栅格数据,压缩方法简单,且能极大地减少存储空间。

游程编码的压缩规则是将同一行(也可按列,但多数按行进行)中具有相同属性值的邻近像元组织在一起,称为一个游程。每个游程用一对数字表达,其中第一个数字为游程长度(所延续的列数),第二个数字为这一游程像元的属性值。每一个图像行的游程分别记录。例如,图2-10c栅格图像的第1行,前3个像元属性为12,接下5个为11。因此,这一行的游程长度编码为:3,12;5,11。同理,第6行的游程编码为:1,12;5,31;2,11。

实践中,当图像(行列数)很大且同属性像元区域多时,例如对很大的分类图像,游程编码的压缩比会很高。为进一步压缩,还可以让游程跨行或跨列。

若相同属性值的像元呈面域分布，还可将游程编码扩展到二维操作，这就是块码或块式编码。块码将图像中具有相同属性值的矩形区域（常取正方形）合在一起记录，用行、列号描述矩形区位置，再加上该区的属性即可。

2 链码

链码或链式编码特别适合于表达线状空间对象（线状地物和多边形边界）的栅格数据，其规则是：每条线一链码，链码由该条线的起始像元的位置（行、列数），线的属性值和表达同条线上像元连接的方向码组成。栅格数据中"线"的像元相连关系只有8种情况，可用8个方向码表现之。图2-13给出一种方向码定义：向东连接为0，向东南和南连接为1和2等，依次类推。

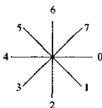


图2-13 一种方向码

现在以图2-11右上部的那条线为例，假定它为高程等于500的等高线；再注意到它的起始像元为第4列第1行，即(4,1)，且起始后，像元向南连接（方向码2），再接着，向东南连接（方向码1），……这样，该条线的链码可编为：500,4,1,2,1,2,1,0,7。

从以上内容可以看到，链码的规则正是基于§2.3.2.2所定义的线的"单一连接"特性。若线在不该中断的地方断了，链码继续不下去；反之，若有三个像元两两间都相互相邻，链码走向失去唯一性，也继续不下去。今后我们还会看到，线的"单一连接"特性是栅格数据向矢量数据转换的途径之一。

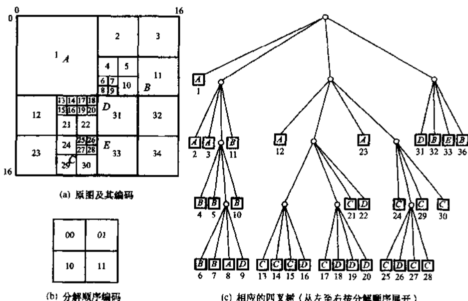
上面看到，游程码和链码对表达多边形和线状对象的图像具有很强的数据压缩能力。那么，压缩格式的数据怎样利用呢？通常有两种途径。一是将压缩数据恢复成完全栅格数据的形式，这个过程称为"解码"或"解压"；另一种是不经解码，直接在压缩状态下操作。不同的压缩编码所能直接进行操作的多少或种类是不同的。例如，线的长度在游程码状态下不易计算。但是，链码却便于直接进行线的长度计算，因为上述方向码若为偶数，意味着线条向东南西北进一格，长度增加1；方向码若为奇数，意味着前进一个像元的对角线，即线长度增加$\sqrt{2}$。还有些操作（如找线的交点等）在压缩格式下一般不如在解码后容易。由此可见，不同的压缩形式各有优缺点。一般而言，如果一种压缩结构不仅压缩效率高，而且在压缩状态下能直接操作的种类多，效率高，就是相对较好的压缩结构。下面的四叉树编码就是这样的一种编码。

3. 四叉树编码

四叉树编码又称为四分树或四象限编码。它要求图像为边长等于2的整数幂（2^n）的正方形（不符合这个条件可通过补零凑齐）。其编码思路是，将图像一级级地等分为四个象限（即4个正方形），无论哪一级哪一个象限，如果经检验是均质的（属性相同）就不再分割，直至整个图像被分为大小不同的若干均质

方块。这样，每一个均质方块就成为很便利的编码单位，因为只要记录每个均质方块在每一级分解中处于哪个方位就行了；而分解方位仅4个：西北（NW）、东北（NE）、西南（SW）和东南（SE）。每个均质方块被分解的级数（或层数），称为分解的深度，深度也体现均质方块的大小。

图2-14给出一个四叉树编码示例，其中图a表示一个属性分别为A、B、C、D和E的分区图像，按上述规则被分解为均质方块；图b定义了一种分解顺序，NW、NE、SW和SE的编码分别为00、01、10和11（二进制表示）；图c给出相应的四叉树，它形似一颗倒栽的树，从上至下、从左至右按分解顺序展开，形象地表现四叉树分解过程及结果。四叉树顶部是根结点，内部结点（分叉点）用圆圈表示；每一个"树枝"的终点用方框表示，称为终结点或叶结点。所有终结点从左至右编号。每一个终结点代表一个上述的均质方块。这样，终结点在树中的位置就能清楚地表达每个均质方块的分解深度，以及在每一级分解中所处的方位。例如，对照图a和c，第21号终结点或均质方块的深度为3（分解到第三层）；在第一级象限中处于西南方位，第二级分解中为东北方位，第三级中又为西南方位等。



(a) 原图及其编码

(b) 分解顺序编码

(c) 相应的四叉树（从左至右按分解顺序展开）

图2-14 四叉树编码示例

四叉树编码按顺序列出每一个终结点（均质方块）的编码，具体编码形式各人不尽相同，图2-15以第9号终结点为例给出 种形式的编码结果。其中，右边4个比特记录终结点的深度，其最大深度可为15（二进制的1111），足以分

解行列数为2^{15}的图像,故一般情况下够用、深度左边的8个比特分别记录第9号终结点在每一层中的方位:在第一级分解中处于东北象限,编码01;第二级中为西南象限,编码10;第三级中仍为西南方位;第四级中为东南象限,编码为11、最左边的比特记录着该终结点的属性值D。

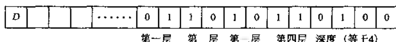


图2-15　第9号终结点的编码

四叉树编码是目前压缩的二维栅格数据结构中较好的一种。它用不同的分辨率表达图像的不同部分,在变化频繁的部分自然地采用高分辨率描述,在图像不变或少变的部分用低分辨率表达。因此,它既可精确表示图结构,又可减少存储量。四叉树编码不仅容易解码,在压缩状态下也能直接实施不少操作。此外,它还可以表达多边形的包含关系,等等。四叉树编码也有缺点。例如,它对图像的要求较苛刻,不仅要为正方形,边长还要等于2^n;若通过补零凑齐,又可能导致分解结果的不稳定:相同形状和大小的图像采用不同的补零方式,可能得出不同的四叉树结构,故不利于形状分析和模式识别。此外,从完全栅格结构数据构建四叉树比较费时,某些操作在四叉树结构下不方便或效率不高等。

4. 地学栅格图像压缩不同于一般图像压缩的特点

由于近年来互联网发展很快,网上图像运作及其压缩的问题很受重视,这里有必要分析一下地学栅格图像压缩不同于一般图像压缩的特点。

第一,地学中有时用栅格图像表达线状地物和分区、分类的面状地物,这种栅格图像中常有像元值连接相同的情况,压缩价值较高。而互联网和多媒体运作的图像大多是照片等实录景观型的栅格图像,其像元数值变动较频繁,很少有像元属性值连片相同的情况,压缩价值不高。不难证明,当像元值变动频繁到一定程度,压缩格式的数据量反而会比原来的完全栅格结构大。例如,像元数值变动最频繁的黑白相间的棋盘样式的图像,若采用四叉树编码,所需空间反而是完全栅格数据的2倍。

第二,为了解决大数据量图像网上有效传输的问题,互联网和多媒体运作的图像常常采用"有损压缩"的办法。其主要思想是,基于认知科学研究,在人眼可分辨的限度内,将图像中数值相近的邻接像元视为相同;或减少像元的亮度级别,例如本来256级亮度简化为16级亮度。经过这样的处理,图像的压缩率自然大幅度提高。实践表明,这种方法实际有效。有时一幅近1 000k字节的BMP(Windows的默认图像格式)图像,压缩为30~50 k的JPG图像,人眼看不出有什么差别。

但是，这种有损压缩一般不适合于地学遥感图像。遥感图像也是实录景观型的栅格图像，像元数值变动通常也较频繁，为什么不采取有损压缩的办法呢？这是因为，遥感图像是供科学分析用，而不是供欣赏的。事实上，当今的图像处理可分为两个大类。其一是科学型的图像处理，它以科学分析为目的，要求"真实"与"精确"，因而不能轻易"有损"。另一种是非科学型的图像处理，它不以真实而以艺术美为准则，例如将扫描的肖像修饰再修饰；或者以商业利益为目标，只要观赏价值基本保持就行。现在互联网上传输的绝大多数图像都是非科学型处理的图像。

由此缘故，地学遥感图像在大多数情况下仍然采用完全栅格结构。若要压缩，也要求进行"无损压缩"。地学遥感图像一般数据量都很大，怎样在"无损"的前提下进行压缩和网上有效传输，是当前一个急需研究解决的重要问题。

此外还应当说明，由于矢量数据结构更适宜表现点、线和多边形空间对象，地学中用栅格图像表达点、线和多边形的情况并不多。在 GIS 中栅格数据的主要应用领域是遥感图像处理和地表三维立体表达，而这两个方面都不很需要栅格数据压缩。因此，除了较深入的 GIS 研究开发者外，一般 GIS 人员和学生接触栅格数据压缩结构的机会并不多。为此，本书上面讲述栅格数据压缩，仅着重基本思想和概念。

2.3.4 栅格数据的采集、输入和输出

1. 栅格数据结构在 GIS 输入和输出上的优势

在地理信息系统的输入输出上，栅格模式较矢量模式占明显优势。其原因有两个方面。一方面，栅格数据结构用数字矩阵来表达，结构简单，数据文件按顺序隐含像元的地址，非常适宜于输入、输出设备的点阵方式实际运作。数字摄影机、数字摄像机、遥感器、数字扫描仪等多种采集设备都生产栅格形式的数据；许多输出设备也是基于栅格模式，如显示器、行式打印机、静电绘图仪、喷墨绘图仪等。

另一方面，栅格图像在地学以外领域，特别是多媒体和互联网领域的应用，比矢量数据要广泛得多。而广泛的应用是科技发展最强有力的驱动因素。20 世纪 90 年代以来，随着互联网、多媒体技术及其应用飞速发展，栅格（点阵）模式的硬件技术进步很快，基于栅格数据结构的输入和输出设备的性能价格比迅速提高。在不到十年期间，大众型的彩色扫描仪降价达数十倍；彩色喷墨绘图仪以数万元一台的价格取代了数十万元一台的静电绘图仪等。因此，现在 GIS 领域无论是输入还是输出，栅格模式的设备占主要地位。

2　栅格数据的采集、输入

地学中栅格结构的数据获取和输入途径主要有下述四种：

第一种是人工方式。人工方式相当于用一张透明薄膜纸蒙在需要数字化的地图上，薄膜纸上绘有规整格网，其外框与地图图幅相同；然后，如图2－10所示的那样，用每一个像元在地图上的对应位置的地理属性作为像元数值，从而可写出如图2－10c的矩阵数组；将此矩阵数组输入计算机，就形成栅格数据文件。

在这个过程中，唯一略需斟酌之处是有线条（线状地物或分区边界）穿过的像元的取值问题。事实上，取值方法没有一定之规，完全视研究目的或应用特点而定。例如，若穿过像元的是分区边界，可以如图2－10那样，采取在该像元中占较大面积的分区代码值；也可采用其他办法，例如按照分区的重要性取值，或根据像元中心在哪个分区取值。若穿过像元的是线状地物，则需注意并不是任何被穿过的像元都能取值。请再参见图2－11，可以发现有不少被线条穿过的像元，如(6,2)、(7,3)、(8,5)等，皆未取值。这是为了保证线状地物的"单一连接"的特性（单一连接的定义见§2.3.2.2）。

第二，从矢量数据直接转换。这一点将在下一节中讲述。

第三种途径是遥感数据（已见本节前面解释）。遥感数据作为电子数据文件，可直接读入计算机中。

第四，扫描数字化。这一点需要稍加说明。

扫描数字化通过数字扫描设备（扫描仪）将硬拷贝（纸张、薄膜等）上的图形图像，转化为栅格数据。GIS的扫描对象一般是地图，因而需要较大幅面的扫描仪，即工程扫描仪。由于用户数量有限，工程扫描仪不可能像现已大众化的小幅面(A4)彩色扫描仪那样大幅度降价，且不谈彩色工程扫描仪，黑白工程扫描仪至今仍比手扶跟踪数字化仪贵很多。常用的工程扫描仪是滚筒式自动光学扫描仪，其主体为一个旋转的滚筒，扫描时将地图固定在滚筒上。当滚筒旋转时，有光源和感应器沿滚筒来回匀速走动，均匀地、依次地感应一个个纸面单元的亮度，并将一行一行的记录直接输入到计算机内，形成扫描栅格图像。目前市场上工程扫描仪的分辨率一般为300～600 dpi(dots per inch，即每英寸扫描感应的点子数)，能满足地图扫描分辨率的要求。

虽然工程扫描仪仍然比手扶跟踪数字化仪贵几倍，但由于扫描矢量化技术的迅速发展，经扫描的栅格地图向矢量数据格式的转换变得愈来愈简便。相比之下，手扶跟踪数字化方式的精度、工作速度和效率愈来愈显不足，其价格优势已逐渐失去意义。由此缘故，扫描矢量化已成为现在硬拷贝地图输入的主流方式。"扫描矢量化"，即将扫描仪形成的栅格地图数据转换成矢量数据格式的原理，将在下一节讲授。

3. 栅格数据的输出

栅格数据的输出是栅格数据每个像元之值转化为输出画面上相应点位处灰度值的过程。我们曾多次说过，栅格数据模型相当于涂抹画的表达方式，栅格图像通过画面上每个单元的明暗或色调来表现空间实体的界限和形态。由于栅格数据结构用数字矩阵来表达，而许多输出设备也是基于栅格模式，栅格结构数据输出时，人们不难利用每一个像元之值来直接或间接地控制输出的强度，使输出画面上相应点位处产生需要的灰度值。

初学者一般很关心彩色输出的问题。现在以计算机荧屏彩色输出为例来简要解释。事实上，任何颜色都可以用一定比例的三原色：红、绿、蓝（RGB）搭配调成；而荧屏上涂有三原色感光材料。为此，栅格数据彩色输出一般需要三个栅格数据文件。当栅格数据彩色输出时，三个栅格文件分别控制计算机显示设备的红、绿、蓝电子枪。

注意，三个栅格数据文件的格网是完全相同的（可参照图2－12理解）；而荧屏也按计算机指令被分为规整格网，例如平常我们说某荧屏显示为800×600的分辨率，就是说该荧屏被分成800列、600行的格子。这样，同一像元在三个栅格文件中的数值，就可分别控制红、绿、蓝电子枪射向显示屏上对应像点的辐射强度。例如，第一行一列的像元在文件1、文件2和文件3中的值，分别控制红、绿、蓝电子枪，在荧屏上第一行一列的像点上，产生所需比例的红、绿、蓝综合叠加效果，形成所需的明暗和颜色。

当输出时只有一个栅格文件，那只能产生黑白图像，因为每个像元只有一个值，该值带动RGB三枪，其综合效果为黑白色。这时栅格像元的属性数据控制相应点位上的灰度，即黑白明暗程度。

在利用三个栅格文件进行彩色输出时，如果三个栅格数据本来就是由红、绿、蓝三原色波段采集来的，而且又分别用它们控制红、绿、蓝三色枪，那么，输出的就应是"原汁原味"的彩色图像。一般民用的摄像、放像设备基本上是如此运作。在遥感图像处理中，卫星影像一般是多波段的，人们常用任意三个波段去控制红、绿、蓝三色枪，输出的彩色经常不是原汁原味的彩色图像。例如，美国陆地卫星TM数据7个波段中，只有第1、2和3波段是通过蓝、绿、红三个可见光波段采集的；用这三个波段的像元数值，去分别控制显示设备的蓝、绿、红三个电子枪，显示屏上出现的大体就是天上看到的地面彩色图像。但是，如果采用其他三个波段的组合，或虽采用第1、2和3波段但不按蓝、绿、红的顺序去控制电子枪，那么，所得到的彩色输出图像，都不会是从天上看到的地面彩色图像，而是有关专业人员常说的"假彩色"图像。

§2.4 基于两种数据结构的进一步讨论

2.4.1 矢量和栅格数据结构的比较分析

1 矢量和栅格数据结构总体比较

矢量结构和栅格结构是利用计算机形象表现客观世界的两种基本方式，二者互相补充，相辅相成。矢量结构与地图渊源深厚，它们力图精确而简明地表达人类对地理空间事物简化抽象的结果，即点、线和多边形空间对象的轮廓，而不关注画面上的其他部分。因此，数据结构具有天然的精练性，以及为了保证准确、精练而带来的结构复杂性。栅格结构的基点是从某种（属性）角度，用简单规整的格网来模拟空间景观的整体形象。所以，它"属性明显，位置隐含"（在行列之中），画面上处处有内容，数据量大而结构简单；但抽象和简明地表达主要空间对象及空间关系的能力不够强。

有趣的是，两种数据结构的表现手法和总体效果也正好相反：矢量数据的空间位置坐标取值可以是任意的、连续的，但表达的空间形象是分立空间对象组成的画面，即总体效果是不连续的；而栅格结构的数据取值方式是不连续的、分立的，但总体表达效果却可以是连续的，表现为照片般的空间图像。

由于与地图的渊源，矢量结构至今在 GIS 领域占主导地位；但栅格结构由于遥感及互联网、多媒体的迅速发展，重要性正在不断增加。

2 矢量和栅格数据结构具体优缺点比较

上述基本点决定了两种数据结构的一系列具体的优缺点。对二者优缺点的一个较全面的分析比较，如表 2-4 所示。表中，矢量结构占优的方面位于上部若干行；栅格数据占优的内容放在表的下面部分。从该表我们可以进一步体会到两种数据结构的互补性。表 2-4 中不少内容已无须解释，仅有几点需要说明。

第一，在数据输出方面，现有的 GIS 书籍都把矢量结构归于优点一方，这是因为早先计算机存储能力弱，图像像元难以精细；且矢量数据通过笔式绘图仪输出在性能价格比上占优。现在情况变了，结论应当修改（参见 §2.3 末尾）。

第二，表中有几条涉及二者在空间分析方面的比较。由于较详细的空间分析内容要到第五章讲解，这里有必要作两点解释。首先，栅格数据适合于空间叠置分析的主要原因，是它在多层因子时采用完全一致的格网（参见 §2.3.2.3 和图 2-12），以致在进行多因子综合分析或叠合分析时，地面单元始终不变而矢量结构不然。参照图 2-12 左边图形就不难理解，矢量数据层叠合分析的结果，会产生很多碎小的地面单元。其次，有些 GIS 书籍上说栅格结构进行数

学模拟容易，而矢量结构难。我们认为不能一概而论，因为矢量结构虽然在模拟数理方程或连续变量场方面不如栅格结构，但是在网络分析的数学模拟和水文及水力学的流管模拟等方面，较栅格结构为优。

表 2-4 矢量和栅格数据结构优缺点比较

比较内容	矢量数据结构	栅格数据结构
面向目标和主要空间对象	较好	不明显
数据结构精练性和数据量	较精练，数据量较小	不够精练，数据量较大
数据精度	较高	不如矢量数据
描述拓扑关系	较易	较难
图形操作	较易	较难
网络和流管分析	较易	较难
比例尺变换、投影变换	较易	相对费时些
形象表现力和实录能力	较弱	较强
数据结构的简易性	结构较复杂	简易
软件开发	较繁难	相对简便些
输入输出	图形精美但愈来愈多借助栅格	输入输出的主要途径
空间叠置分析等	较难	较易
模拟连续变量场	较难	较易
数据共享的便利性	较差	较好
三维立体形象表达能力	较差	较强

第三，矢量结构软件开发较难，是指按国际上矢量 GIS 软件和栅格 GIS 软件的现有模式衡量，开发矢量 GIS 软件的难度较大，编程较为繁难。

第四，在信息共享方面，目前还没有一种公认容易共享的矢量数据格式；而栅格数据，尤其是完全栅格数据结构易于为大多数人理解和使用。

第五，栅格数据适合于三维立体形象表达，这个道理就如同油画比线条画的立体表达能力强一样。平面上的线条画也能通过透视关系来体现立体形象，例如勾画远山的轮廓。但是，油画不仅能够通过明暗、色调差别来表现轮廓，还能表现光照下的阴影和质地。栅格数据结构就完全具有油画的这种形象表现能力（详见 §4.4.3）。

3. 矢量和栅格数据结构的发展

20 世纪 80 年代末以前，矢量和栅格数据结构基本上是分别发展的，形成了那时 GIS 技术的两大派别：矢量型 GIS（vector GIS）和栅格型 GIS（raster GIS）。

矢量型 GIS 一直被认为是 GIS 主流技术。矢量型 GIS 来源于早期的 CAD 技术。早期 CAD 多基于 Unix 和 Vax 系统，其中一些 CAD 软件业已发展了空间数据和地理属性数据处理的功能。这些早期 CAD 技术与后来以 AutoCAD 为代表的，主要基于 PC 而迅速流行起来的

CAD技术不尽相同,后来的CAD致力于发展设计制图功能,其"属性"仅是为绘图和显示服务的图形特征。矢量型GIS在早期CAD技术的基础上,进一步发展了空间数据、属性数据管理和空间分析的功能。到80年代,矢量型GIS借助于迅速发展和成熟起来的关系数据库技术管理地理属性数据,并通过公共标识符使点、线、面图形与地理属性数据一一对应地相挂联,发展成为集空间数据采集、显示、贮存、处理、分析、建模等功能模块于一体的大型软件包,形成"集成式GIS"技术。

矢量型GIS技术的典型代表即美国ESRI公司研制的软件Arc/Info,80年代中前期该公司将其著名的地学关系数据模型付诸实践,首先成功地开发出集成式GIS技术。稍后的Intergraph公司的MGE也是矢量型GIS软件的重要代表。MGE软件在microstation(一种CAD软件)的基础上,开发出一系列模块来实现GIS功能。MGE完全采用商用数据库,如Oracle等,来管理地理属性数据,这与Arc/Info不同,Arc/Info除了能挂接商用数据库来管理属性数据外,自身还带有一个不大的内置数据库(Info)。

栅格型GIS最初产生于进行资源与环境(特别是土地资源)评价的计算机模型。事实上,加拿大的世界首例GIS就是一个栅格型GIS。该系统最初采取栅格数据格式,固然有计算机能力限制的因素,但栅格型GIS确实适宜于多因子叠合分析,也是一个重要的原因。栅格型GIS的发展更多地得益于遥感技术之进步。从20世纪70年代开始,美国陆地卫星MSS和TM图像在资源环境众多领域得到广泛的应用。为了从多波段的遥感图像提取所需信息,栅格型GIS软件伴随着遥感图像处理技术的进步而不断发展。由于多波段分析实质上也是一种多因子综合分析,遥感图像处理技术的发展极大地丰富了多因子自动综合分析的手段和内容。栅格型GIS软件的代表是ERDAS。ERDAS是一个集成化的大型遥感图像处理软件,它也自称是一个栅格型GIS软件,其相当多的遥感图像处理功能在软件中被用来作为多因子分析的手段。在ERDAS的技术总说明书(Field Guide)中,有专门的章节讲述栅格型GIS的概念,以及介绍如何应用ERDAS软件功能去实施栅格型GIS项目。

由于计算机及GIS发展水平的限制,在很长时期内,矢量型GIS和栅格型GIS之间很难相互转换;二者间的鸿沟致使两种技术派别的争论一直在持续。直到20世纪80年代后期,两种格式之间的转换技术成熟,矢量型GIS和栅格型GIS才成为一体化技术。由于矢量型GIS一直被认为是GIS主流,这种一体化也通常称为GIS技术与遥感技术一体化。这是GIS技术基本成熟的重要标志。

顺便说明,几年以后,即20世纪90年代前期,新兴的全球定位系统(GPS)技术的性能价格比达到普及应用的程度,GIS技术、遥感(RS)技术与GPS技术三者一体化,形成3S技术(回顾§1.3.3.2)。

4. 矢量和栅格数据结构的选择

栅格GIS与矢量GIS之一体化,并不改变两种数据结构各有所长的事实。因此,在GIS实践中,必须在两种数据结构中作出恰当选择。在更好的,甚至大一统的GIS数据结构(如果能形成的话)出现以前,建设GIS实用系统的较好方案是对两种数据结构兼收并蓄,同时根据实际应用的需要与可能,来决定具体的应用方式和内容。

在大多数情况下,GIS应用部门需要应用地图及其查询分析。因此,矢量数据结构通常是不可缺少的。但是,有些用户还同时需要进行地表景观的三维显示和分析,那就必须加上应用栅格数据结构。还有些用户,特别是进行较大范围较小比例尺的自然资源、环境、农业、林业、地质等区域问题研究的用户,常常需要采用遥感数据,那就更需要将栅格数据结构与矢量结构相结合。他们常常将遥感图像处理所提取的分类分区和线状地物等信息,从栅格形式转换到矢量结构的软件平台上,绘制矢量地图。

采用什么数据结构还要视条件而定。例如,1999年出现的美国1m分辨率的卫星数据对很多应用领域都很有用,但是它比较贵,超出国内许多单位愿承受的范围。总之,具体情况具体分析,尽可能在许可的条件下发挥矢量结构和栅格结构的长处。

2.4.2 栅格、矢量数据结构的相互转换

1. 矢量数据向栅格数据转换的主要步骤

矢量数据向栅格数据的转换,称为矢量栅格化(rasterization)。我们在§2.3.4.2曾讲过人工方式获取栅格数据的方法,即在需要数字化的地图上,蒙上范围与地图相同的规整格网的透明薄膜,然后根据每一个格子所对应的地图上的内容,来确定该网格的属性值。矢量栅格化的转换思路与此非常类似。具体转换过程可以概括为如下几个步骤:

第一步,准备好矢量数据或矢量地图。

第二步,构架一个与地图等大小的格网,根据精度需要与条件许可,选取好适当的格网密度或分辨率。设所选正方形像元边长为Δ,则:

格网的列数 = 地图图幅长 / Δ

格网的行数 = 地图图幅宽 / Δ

根据此行、列数在计算机程序中定义一个数组变量。如果该地图是可以借以建立地理坐标的(例如地形图),则根据该地图确定第一行第一列像元的地理坐标值(x_0, y_0),以及用地理坐标值表示的每个像元的宽度Δx(参见§2.3.1.2中公式)。

第三步,根据地图上的点、线、多边形相对这个格网的配置及其属性来确定数组变量每一个像元的数值。

最后,将上述矩阵数组结果存储于某种所需的图像格式中,从而形成可以在相应软件平台上显示的栅格图像。

上述步骤中,第三步内容较多,需要专门说明。

2. 点状、线状和面状空间对象矢量数据的栅格化

怎样根据地图上的空间对象及其属性相对于这个格网的配置来确定数组

变量每一个像元的数值呢？下面分别就点、线和多边形对象加以说明。

第一，点状对象的栅格化。点状对象的栅格化十分简单：该点落入哪个像元，就根据该点状对象的特性赋予该像元属性值。

第二，线状对象的栅格化，也称为弧段栅格化。线状对象在矢量数据中，除属性外，是一连串(x,y)值。弧段栅格化要求对每两个相邻的(x,y)之间的线段所穿过的像元赋值；但必须保证赋值的任何线状地物像元之间在北、东北、东，……西北8个方位上的单一连接关系（"单一连接"的定义参见§2.3.2.2和图2－13链码有关内容）。为此，不可能将所有被线段穿过的像元都赋值。参见图2－11，以左下方的那条线为例。为了保证该线赋值像元间的单一连接性，一部分像元，例如第2列第6行的像元，即像元(2,6)，虽然有小段矢量线穿过，但不能栅格化，仍然保持零值；否则，该像元就与像元(2,5)和(3,6)构成了皆"两两相邻"的三个像元，不符合"单一连接"的定义。同理，像元(3,7)、(5,7)和(8,8)也都有小段矢量线穿过，但皆保持零值。

保持单一连接性的具体方法不止一种，这里不进一步讨论。

第三，多边形对象的栅格化。多边形对象的栅格化又称为多边形填充，它要进行两方面工作。其一是决定多边形边界所穿过的像元应归于那一边的多边形（分区）。这一点在§2.3.4.2的人工采集栅格数据内容中已说明过，即可以取在该格子中占较大面积的分区的代码值（参照图2－10），但也可按照分区的重要性取值，或根据像元中心在哪个分区来取值等。

另一件工作是自动探测哪些像元位于哪个多边形的内部，以便根据该多边形的特性给这些像元赋属性值。这就是"多边形填充"的真正含义。多边形填充也有多种算法。例如，如果原来的矢量数据中有多边形标签点（参见§2.2.2）数据，则可在多边形边界穿过的像元都确定后，利用该多边形的标签点给像元赋值，并向四周扩展，将遇到的未赋值像元皆赋予该多边形属性，直到遇上已赋值的该多边形边界为止。多边形填充还有若干其他方法。这里就不一一介绍了。

矢量数据栅格化的最重要应用之一，就是前面所说的矢量数据输出，使矢量地图得以通过绘图仪、打印机等栅格（点阵）模式的输出设备输出。

3. 栅格格式的遥感分类图像向矢量数据的转换

栅格数据向矢量数据转换也称为栅格矢量化（vectorization）。不同类型的栅格数据（参见§2.3.2.2）矢量化的方法不尽相同。目前应用较多的栅格矢量化主要有二种形式：一是将遥感图像处理所提取的栅格格式的分类图像转换为矢量结构；二是将扫描仪扫描地图所获取的栅格数据转为矢量数据形式（从遥感图像中提取的线状地物的矢量化亦可归于此类）。下面我们将沿着这两条线索来讲解，本小节先讲前者，下一小节再讲后一种形式。这两种形式大体上能覆盖GIS栅格矢量化的主要内容。

在§2.3 2.2中我们说过，遥感图像处理信息提取的主要任务，是从实录景观的遥感图像中区分、提取出具有相同性质的区域（面状地物），以及线状地物和点状地物。其中，从某种学科角度对遥感图像进行分类，是最常见的工作。当遥感图像被区分为不同地类或不同植被等分区时，就形成如图2-10所示的那种的"分类图像"。从图2-10可见，分类图像是相当简单的，只要将各分区的边缘线提取出来，就形成了矢量数据。例如在图2-10c中，属性值为11、12和31的区域的格网分界线，是容易提取出来的。但是，如此提取的边缘线是折线形态的（参见图2-16左部），还必须进行§2.2 4所讲述的矢量编辑工作。首先，要进行圆滑，增补一些中间点，同时去除或移动一些不必要的中间点，使折线变为弧线；然后，再进行各种必要的图形编辑、属性赋值、拓扑编辑等，最后才能形成完好的矢量数据（参见图2-16下部）。

4. 扫描地图的栅格图像向矢量数据的转换

扫描地图的栅格图像向矢量数据的转换要相对麻烦一点。由于地图上线条本身宽度通常比扫描仪感应像点还大，扫描仪扫出的地图线条常常有不止一个像元的宽度，以致放大显示扫描图时，会看到很多像多列行军队伍般的宽线条；同时，地图上原"空白"的地方可能因为褶皱或脏物，扫描后像元值也常常并不等于零。遥感图像中提取出的线状地物也常常是多列像元的宽线条。对这样的情形，需要进行下列步骤的工作：

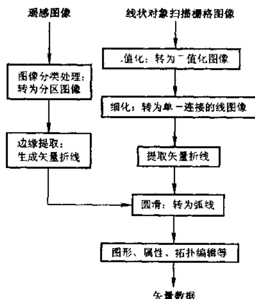


图2-16　栅格数据向矢量数据的转换步骤

第一步，进行二值化，即通过高通滤波等图像处理办法，将栅格图像变成二值化图像。注意，由于地图上各图元的颜色和深浅各不相同，扫描图像中的点和线会有各级灰度值。所谓“二值化”，就是把扫描图像中点和线的像元属性变更为1，使整个图像只有0和1两种取值（参见图2－16右部）。

第二步，将多列行军队伍般的宽线条（这时属性皆为1），转变成图2－11中可见的“单列队伍”的线条，即具有“单一连接”特性的线条（单一连接的定义见§2.3.2.2），这里面既有“剥皮”，即把两边像元剥去的工作，也有一旦线被“剥”断时的断线连接工作。这一步骤称为“细化”。从图2－11可见，从单一连接的线条很容易提取矢量线：或者取线条的某一边的边界，或者提取线上每个像元中心的连线。当然，这样提取出来的矢量线也都是折线，也需进行矢量编辑工作。这些工作前一小节刚讲过，就不重复了。

在实践中，扫描矢量化通常给用户两种选择，即自动矢量化和交互式矢量化。自动矢量化大体就是图2－16给出的思路，常应用于分版的等高线图、水系图、道路网等比较清晰的，以线状对象为主的地图矢量化。但在很多情况下，如各种城市大比例尺图的矢量化，一般还得采用交互式。交互式矢量化让作业员采取人机交互的方式，对逐条线划具体处理，包括去除褶皱或脏物导致的像元，以及不需要的原图符号注记所造成的像元。当然，当线划状态比较好时，应尽可能多地自动跟踪和运作，以减少作业员进行人工干预的工作量。

2.4.3 其他空间数据结构

除矢量和栅格数据外，GIS学者还进行了很多空间数据表达的研究，提出过多种数据结构。虽然其中多数还没能应用于实践，但它们丰富了地理信息科学的内容。由于所提出的数据结构一般都可以看做是二维矢量和栅格两种基本格式的过渡或扩展，因而都放在本节中讲述。一种扩展是在二维范畴内兼具矢量和栅格特点的模型或结构，例如镶嵌数据结构、矢栅混合数据结构等；另一种是从二维向三维扩展，发展三维空间数据模型和结构。对本课程而言，除不规则三角网数据结构外，其他数据结构皆属介绍性内容，不做基本要求。欲多一些了解这方面研究的读者，可参见龚建雅《地理信息系统基础》第四章

1. 镶嵌数据结构

镶嵌（tessellations）数据结构采用各种形状的镶嵌地面单元来拟合地表。同栅格结构一样，它用一个格网来划分整个研究区域，但是，格网的每个网格的形状可以是多种多样的。按网格形状的不同，镶嵌数据结构可划分为规则镶嵌数据结构和不规则镶嵌数据结构两大类。规则镶嵌数据结构的网格采取统一形状的网格单元，且这样的单元能相互镶嵌填满整个地面，如同各种规则单元镶嵌的人行道或地板一样。这样的镶嵌单元有：等边三角形、等腰直角三角形、

矩形、正六边形、平行四边形等。注意正多边形不一定都能相互镶嵌填合地面。不过至今为止，除正方形外的基于规整地面单元的数据结构均不够理想，离应用还有相当距离。

不同学者对栅格数据结构概念的理解不尽相同。有些学者认为，基于正方形的镶嵌数据结构不完全等同于栅格数据结构，也有的学者把栅格数据结构看做是规则镶嵌数据结构的一个特例。

不规则镶嵌结构的地面单元采取不规则的网格形状。最重要的且有应用价值的不规则镶嵌结构是不规则三角网（triangulated irregular network，简称 TIN）和泰森（Thiessen）多边形。参见图 2－17，不规则三角网由不规则分布的点（图 2－17 中的黑点所示）连接而成；泰森多边形由不规则三角网的每条边的垂直平分线构成。泰森多边形和不规则三角网二者为对偶结构。不规则三角网也可以反过来由泰森多边形构造。

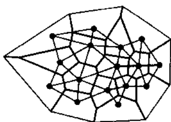


图 2－17 不规则三角网和泰森多边形

在地球科学及其应用领域，由于各种条件的限制，数据采集的样点在大多数情况下是离散的、分布不规则的。不规则三角网可基于离散数据样点直接构造，即直接采用不规则样点构成的三角形作为空间分析的地面单元。同规整的栅格结构相比，TIN 的不规则性会带来某些数据处理的代价，但是，直接或比较直接地应用样点数据带来的好处却常常更有价值。除了保持原始数据精度和可靠性的优点外，不规则三角网还可以减少数据冗余。后者是因为，TIN 的网眼结构本身适应于数据的实际分布：在空间数据或事件密度较高的区域，TIN 的地面单元即分析单元自然地小而密；反之，在空间数据或事件密度较低，因而分析可能允许粗略一些的地方，TIN 的网格单元自然地变得大而疏。栅格数据就没有这种分辨率内部可变的优点，同一个栅格格网的格子或分辨率是固定不变的。由此缘故，不规则镶嵌结构在某些情况下具有比规整结构更好的研究和应用价值。

不规则三角网在 GIS 中最主要的应用是在数字高程模型或地形分析领域。在实践中，通过高程测量可以得到不规则样点(x,y)的高程 z 数据，这时，TIN 的每个三角形本身就是一个有高度、坡度等地形要素，可以构架地表立体表达的地面单元。地形分析的 TIN 数据也常常由地图等高线数字化得到。例如，用数字化仪手扶跟踪一条等高线，实际上就是沿着该线，随机采集具有该线之高程 z 值的一串 x,y 坐标点（参见 §2.2.4）；跟踪另一等高线又会得到另一串具有某个 z 值的 x,y 坐标点。如此采集的 TIN 数据对表达地形有利，因为需要较细致

描述地形的区域，通常就是等高线较密的地方；而这些地方又通常正是数字化采集点和 TIN 的三角形较密集的区域。因此，TIN 数据表示不连续对象，如悬崖、断层、海岸线、山谷谷底等具有优势。TIN 的拓扑结构易于存储，操作也比较便利，例如可以方便地进行地表坡向、坡度和地形特征线，以及自动生成等高线等计算。

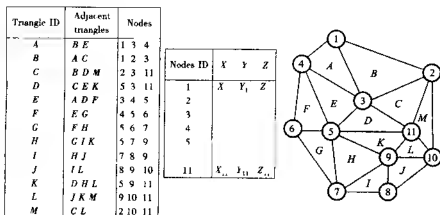
由于 TIN 的上述长处，下面将专门用一个小节介绍 TIN 数据结构的特点。

泰森多边形由不规则三角网每条边的垂直平分线构成。因此，区域中任何一点总是距自己所在的多边形内的采样点最近。这个特点可以很有效地应用于许多场合，诸如数据内插、邻接或接近度分析、可达性分析、最近点和最小封闭圆问题等。

尽管不规则镶嵌结构在某些方面有优势，但对于其他一些空间数据处理和分析任务却无能为力。例如，生成不规则网格比较复杂和费时；两个不规则网格覆盖和叠加分析相当困难等。由此缘故，不规则镶嵌结构的应用还比较有限，需要进一步研究。

2. 不规则三角网数据结构

由于 TIN 对三维立体表达的重要性，有必要具体说明其数据组织方法，以帮助理解。



Triangle ID	Adjacent triangles	Nodes
A	B E	1 3 4
B	A C	1 2 3
C	B D M	2 3 11
D	C E K	5 3 11
E	A D F	3 4 5
F	E G	4 5 6
G	F H	5 6 7
H	G I K	5 7 9
I	H J	7 8 9
J	I L	8 9 10
K	D H L	5 9 11
L	J K M	9 10 11
M	C L	2 10 11

Nodes ID	X	Y	Z
1	X	Y_1	Z
2			
3			
4			
5			
11	X_{11}	Y_{11}	Z_{11}

图 2-18 TIN 数据结构

不规则三角网有多种数据组织和存储方法。最常用的方式是把三角形作为一个基本的空间对象，并将它与相邻的三角形和顶点进行拓扑连接。另一种方法是把三角形顶点作为基本的空间对象，将它与其他顶点相连接。这里介绍前一种方法。假定有一个如图 2-18 右图所示的三角网，为有效组织 TIN 数据，可建立两个表。一个表（图中左表）记录每个三角形的标识符、相邻三角形标识符和它的三个结点的编号；另一个表（图中右表）记录每个结点的坐标值。

这样，三角形及其相互关系，以及三角形与结点关系的查找，就可以容易地进行。

综上所述，TIN 结构用一个格网覆盖整个地区，并试图用格网的每个单元及其属性（坡度、坡向、明暗色调等）组合起来表达地表形态。在这一方面，TIN 结构与栅格数据结构思路相同。另一方面，TIN 的地面单元不是方形，而是不规则的三角形，它不能像栅格数据结构那样采取方便计算机运作的矩阵形式，因而又不得不借助于矢量数据的方法来描述三角形地面单元的点、线及其拓扑关系。由此可见，TIN 数据结构兼具栅格和矢量两种数据结构的部分特点，从某种意义上说，它是介于栅格和矢量两种数据结构中间，具有过渡色彩的一种数据结构。

不规则三角网结构和泰森多边形结构又称为 Delaunay 三角网和 Voronoi 图。Voronoi 图因俄国数学家 Voronoi 于 1908 年发现此几何构造而命名。Voronoi 图由于空间划分上的等分性特征，在不少领域得到应用。在地学界最先应用 Voronoi 图的是气象学家泰森（Thiessen），他在研究随机分布的气象观测站时，按 Voronoi 图对每个观测点划分了多边形范围。因此，Voronoi 图也被称为泰森多边形。在生物学领域，Voronoi 图被称为 Winger Seitz 单元或 Blum 变换。Delaunay 三角网是苏联数学家 B Delaunay 于 1934 年发现的。Delaunay 三角网是 Voronoi 图的对偶，将 Voronoi 图中各多边形单元的内点连接后得到一个布满整个区域而又互不重叠的较优三角网结构

Voronoi 多边形是一种重要的混合结构，融图论与几何问题求解为一体，是矢量、栅格空间模型的中间途径。如果把空间邻接定义为多边形邻接，并把围绕各个物体的 Voronoi 多边形的边界用等距离准则来确定，则所有地图上的物体（此处为点和线段）就具有明确邻居从这个思想出发，就可导出一种统一的途径来处理许多空间问题。

Voronoi 多边形在很多学科中都是一种重要的几何构造，很多几何问题可用 Voronoi 多边形得出有效的、精致的，在某种程度上还可以说是最佳的解。在二维空间，Voronoi 多边形在求解"全部最近邻居"问题、构造凸壳、构造最小扩展树以及求解"最大空圆"等问题中，被用作优化算法的第一个步骤。在模式识别中，Voronoi 多边形的应用也越来越广泛。Voronoi 多边形的建立也是计算两个平面图形集合之间最小距离优化算法的预处理步骤。Voronoi 多边形在地理学、气象学、结晶学（晶体生长模型）、天文学、生物化学、材料科学、物理化学、考古学、生态学等领域均得到应用

3. 栅格矢量混合结构

栅格矢量混合结构可分为两种情况。一种情况是两种数据结构在同一 GIS 系统中混合应用，虽然严格说这谈不上是一种"混合结构"，但这种情形非常普遍，具有应用意义。在这种应用中，矢量和栅格结构的数据同时存储在 GIS 系统中，需要时将二者调入到内存，进行统一的显示、查询和分析。不过，不同系统处理矢量与栅格结合的方式不尽相同。比较高级的方式是能进行矢量栅格的联合查询与分析。

另一种情况是尝试构建有实在意义的一体化栅格矢量混合结构，使之同时兼有矢量和栅格结构的主要优点。

在这一方面，特别值得一提的是我国学者提出的矢量栅格合一的一体化数据结构（李德仁，1990；龚健雅，1993）。在这种数据结构中，既有矢量实体的理念，又有栅格覆盖的思想，力图同时覆盖栅格矢量两种数据结构的特点。该数据结构的理论基础是多级格网方法、三个基本约定和线性四叉树编码。多级格网方法是将栅格划分成多级格网：粗格网、基本格网和细分格网。粗格网用于建立空间索引，基本格网的大小与通常栅格划分的原则一致。但基本栅格的分辨率不足以满足矢量表达的精度要求。因此，点和线（含多边形边界）所经过的基本栅格，被进一步细分为 256×256 或 16×16 个格网，以达到矢量的精度要求。粗格网、基本格网和细分格网都采用线性四叉树编码方法。

为了使上述栅格数据具有矢量特点，该结构的作者提出了点状地物、线状地物和面状地物的三个约定。点状地物仅有空间位置没有形状和面积，在计算机内仅有一个位置数据；线状地物有形状，没有面积，在计算机内需要由一组栅格元子填满的路径表达；面状地物有形状和面积，在计算机内由一组填满路径的元子表达的边界线和内部区域组成。采取措施使每个实体有唯一的标识号，它可以连接属性甚至空间拓扑关系，所以它具有完全的矢量特性；与此同时，由于用栅格元子充填了线性目标和多边形边界的路径以及多边形的内部空间，它又是栅格化的，可以进行各种栅格操作。关于矢量栅格一体化数据结构的详细内容请参阅龚健雅《整体 GIS 的数据组织与处理方法》一书（1993）。

4 三维空间数据结构

上面讨论的数据结构都属于二维范畴。本小节介绍两种三维空间数据结构，它们分别是基于二维矢量结构和二维栅格结构向三维扩展的结果。

一种三维空间数据结构称为八叉树数据结构。

八叉树数据结构是由四叉树栅格结构向三维扩展，应用于矿体等真三维现象的一种三维空间数据结构。对于矿体等真三维目标而言，由于矿物的类型、品位、容重等随着三维空间位置不同而变化，因而表达矿体信息的空间单元必须是三维的，具有 (x,y,z) 三维位置坐标。为了适应矿产储量计算和矿山开采，通常将矿体划分成三维的栅格格网，网中每一个单元为小正方体，通称体元，每一个体元都有一个或多个对应的属性数据。

由于三维栅格需要占用比二维栅格大得多的存储空间，近年来，一些学者在四叉树基础上提出了用八叉树表示矿体的三维结构。八叉树的表达方法与四叉树类似（参见§2.3.3.3），但它不是一分为四，而是将正方体区域一级级地一分为八，无论哪一级的正方体如果经检验是均质的（属性相同）就不再分割，直至整个图像被分为大小不同的若干均质方块；以每一个均质方块作为编码单位，这样就可以得到一棵八叉树。

另一种空间数据结构是基于二维矢量结构扩展而成的三维矢量数据结构。二维矢量结构利用平面上的几何点、线和多边形来描述地球表面的点状、

线状和面状(区域)空间对象;而三维矢量数据结构则加以扩展,试图用三维空间中的几何点、三维曲线、曲面和三维多边体,来描述真实的三维地理空间中的地理事物和现象。但由于非常繁难,三维矢量数据结构研究进展至今很小。

5. 简短的小结

尽管在地理空间数据结构上进行了多种多样的努力,但由于简便易行的平面图画与人类视觉欣赏习惯的深厚渊源关系,以及构架三维空间数据结构的繁难性,作者认为,在今后相当长时间内,二维(含2.5维)数据结构仍将是地理空间表达的主要手段。作者还认为,尽管已有很多努力试图调和、兼容或集成矢量与栅格两种数据结构的特点,代表二维形象表达两种基本思路和手段的矢量和栅格两种数据结构的特色和魅力,将是永存的。

尽管构架真三维空间数据结构有很多困难,但一旦成功,将对人类思维方式和研究手段带来革命性影响。这是有重要意义的GIS前沿研究课题之一。

在GIS实践中,要善于综合应用现有的数据结构。例如,一个采集某种属性的样点集可以采取多种表达手段,如可以采用带空间坐标的属性表的形式;可以形成不规则三角网数据;还可以通过数据内插转换为基于规整格网的栅格数据;最后还可以再转换为等高线等矢量数据。总之,具体情况具体分析,以尽可能博采众家之长为宗旨。

第三章
地理空间数据管理

地理空间数据管理主要指地理空间数据库(或 GIS 数据库)系统及其管理。第一章说过,数据库系统是所有信息系统的基础;地理信息系统自不例外。本章与上章的关系是,二者可以说都从属于"地理空间数据组织管理"这个大题目;但上章主要说明地理数据怎样在计算机中组织,以表达地理空间事物和现象,所涉及的数据组织工作及有关计算机运作属比较"基层"的范畴,谈不上数据"管理";而本章主要讲述怎样组织实际的,能有效管理的数据集合,以便更好为用户服务。如果说上章的内容相当于解说一般书籍和文献产生的来龙去脉的话,那么,本章的内容就相当于讲解怎样构架一个个图书馆,并建立好管理机构,以便更好地为读者服务。

数据库系统是计算机科学中一个很重要的分支,内容非常丰富,且发展很快。但是,由于第一章所述的地理空间数据的特殊复杂性,一般计算机科学的数据库系统技术和理论还不能很好解决 GIS 数据库的问题。由此缘故,我们将一般的计算机数据库称为常规数据库或通用数据库。GIS 数据库与通用数据库的关系是:一方面,前者基于后者的理论或理念,并直接应用后者的技术;另一方面,GIS 数据库又常常局部地直接基于操作系统及其文件管理系统,来解决通用数据库难以处理的特殊问题,从而从特定角度对通用数据库系统的技术和理论作出贡献。为此,本章第一节首先介绍一般的数据库系统和计算机文件管理系统的基础知识;在此基础上,第二节再讨论地理空间数据库的特点和拓展,以及它怎样应用通用的数据库技术;最后在第三节讲述近年来很受重视的,为管理地理空间数据服务的元数据问题。

§3.1 数据管理基础知识

本节讲解数据管理基础知识。为此用一整节的篇幅(不少 GIS 书籍不这样做),是因为作者感到这部分知识不仅对理解 GIS 基本构架十分重要,而且也是理解第七章中某些 GIS 新技术内容的基础。

3.1.1 数据库及其管理系统的基本概念

1. 数据组织和管理的层次

数据库中数据组织和管理是分层次的。如同一种期刊的结构一样，从物理上可以划分为“字、行、页、期和卷”等层次；从逻辑内容上，可分为“词、文句、段、目、节和章”等层次。同样，数据组织和管理的层次通常也有两种划分：从物理上，即从计算机物理存储的角度，数据组织管理的层次可分为：比特、字节、字、块（物理记录）、桶和卷；从逻辑内容上，即从数据与其所描述之对象间关系的角度，或从应用角度，数据库中数据组织和管理的主要层次是：数据项、记录、文件和数据库。后一种层次划分是本小节主要关注的对象。下面分别解释之。

● 数据项。数据项是可以定义数据的最小单位，也称为字段等。数据项与现实世界实体的属性相对应，它可以采取数值、字母、字母数字、汉字等形式。数据项有一定的取值范围。数据项的物理特点在于它具有确定的物理长度（一般用字节数表示）。多个数据项可以组合，构成组合数据项。如“日期”可以由日、月、年三个数据项组合而成。

● 记录。记录与现实世界的实体相对应，由若干相关联的数据项组成，每个数据项表示实体的某种属性。换言之，记录是关于一个实体的数据总和。例如，某单位的员工数据库中，每个人（实体）有一条记录，其姓名、性别、年龄等就是构成了该人记录的数据项。为了检索数据库中的记录，每个记录必须有唯一标识符，称为主码或关键字，唯一标识符常常是记录的序号，且放在第一个数据项的位置。

● 文件。文件是某种类型逻辑记录的集合。例如，一幅图的点状地物数据可能构成一个数据文件，一张属性数据表也可以是一个文件。文件用文件名标识。文件根据记录的组织方式和存取方法可以分为：顺序文件、索引文件、直接文件、倒排文件等。

● 数据库。数据库是比文件更高层次的、真正实施数据管理的数据组织。下面具体讨论之。

2. 数据库

数据库是数据管理的高级阶段。作为文件之上的更高层次的数据组织，数据库可能包含成千上万个的文件，例如，一个 GIS 数据库可能含有数百、数千幅图，每幅图又可有点、线、面多种数据文件和多种属性表。但是，数据库并不是众多文件的简单集合，数据库中的文件及其内的数据是有内在联系的；数据库系统透过文件，全面有效地实施数据管理。

数据库的具体定义在不同文献上不尽相同，但大体可分为两类。一类从某种应用的角度，定义数据库为被存储起来的数据集合，这些数据被特定的组织，

如公司、大学或政府机关等所应用。另一类定义注重从技术本身特点来描述。简捷的如:数据库是以特定结构组织和存储,以便有效管理的相互关联的数据及数据文件的集合。较全面的如:数据库是长期储存在计算机内的,有组织的可共享的数据集合;数据库中的数据按一定的数据模型组织、描述和储存,具有较小的冗余度、较高的数据独立性和易扩展性,并可为各种用户所共享(萨师煊等,2000)

数据库的职能与图书馆类似。为帮助理解数据库概念,让我们以图书馆为对照来说明上述萨师煊定义中的主要成分(表3-1)。首先,数据库的数据相当于图书馆中的书籍。图书馆的图书是供众多读者借阅利用的;而数据库的数据也要能方便众多用户存取、分析和共享。数据库中的数据按一定的数据模型组织、描述和储存;图书馆的图书、书架也要按一定的编目方法来放置。图书馆中的图书不依赖于读者和借阅方式而存在,数据库中的数据也有较高的独立性。图书馆的藏书购书从来讲究精简;同样,数据库也要求减少冗余。最后,也是最重要的,正如图书馆的运作和功能靠图书馆管理机构和人员来实施和实现一样,数据库也需要自己的“管理机构和人员”,这就是“数据库管理系统”。

表3-1 数据库与图书馆之比较

数据库	图书馆
数据	图书
数据检索	图书查询
文件	按类存放的一组图书
数据库用户	读者
数据存取分析	图书借阅利用
数据所占用的外存	书库
数据模型	编目的方式方法
数据的物理组织	藏书间和书架的安置
数据库管理系统	图书馆管理机构和人员

3. 数据库管理系统

数据库管理系统(database management system,简称DBMS),是进行数据库存取和各种管理控制的软件,是数据库系统的中心枢纽,用户(及其应用程序)对数据库的操作全部通过DBMS进行。通常说的数据库系统软件平台主要就是指DBMS软件,例如当前常用的大型数据库软件Oracle和SQL Server,以及小型数据库软件Visual FoxPro和Access等。

数据库管理系统的功能因不同的系统而有所差异,但一般都具有以下主要功能:

• 数据库定义功能。给用户提供定义概念模型及相应数据库要素的能力。定义数据库即创建数据库框架，为库内填充数据准备好条件。

• 数据库管理功能。包括整个数据库的运行控制、数据存取、更新管理、数据完整性及有效性控制，以及数据共享时的并发控制等。

• 数据库维护功能。主要有数据库重新定义、数据重新组织、性能监督和分析、数据库整理、发生故障时恢复运行等。

• 数据库通讯功能。包括与操作系统的接口处理、与各种程序语言的接口，以及与远程操作的接口处理等。

在数据库建设的实践中，首先需要对应用项目的实际情况进行系统分析和设计；同时选择好适合的 DBMS 软件平台。在此基础上，用户可以通过该平台上的应用程序，开始创建或定义数据库。数据库一经创建或定义，用户就可以往数据库中输入数据。例如，将每个学生的学号、姓名、年龄、性别、所在系等数据，输入学生情况表等，这就相当于图书馆设施和管理条件已具备，可以往书架上放置图书一样。数据库的建设通常是一个不断充实、应用和修改的过程，其间要进行大量的数据选择、查询、插入、删除、修改等操作。

数据库建设的分析设计一般包括三项内容。一是进行需求分析，写出系统需求分析报告；第二是对欲建库的数据及语义进行详细描述，对数据约束和数据之间的关联进行详细描述，写出数据库信息要求报告；三是详细描述数据库的数据操作要求、处理方法和处理流程，画出系统功能模块图，完成数据库的操作和应用要求报告。创建或定义数据库的工作（若采用关系数据库），主要包括：创建和查看数据库属性、创建表、确定表的关键字和约束条件，以及查看和修改表结构等。

3.1.2 数据管理技术的发展：文件系统与数据库系统

为了帮助理解数据库概念及其"数据独立"、"有组织"和"可共享"等特点，有必要回顾一下数据管理技术的发展过程。这样做还有一个目的，即顺便讲解文件管理系统。了解文件系统也是学习 GIS 的需要，因为如前所述，GIS 与一般管理信息系统不同，一般管理系统如学籍管理系统或材料管理系统等，基于通用数据库系统软件平台就能建设；但 GIS 的某些特殊要求通用数据库难以满足，GIS 数据库系统常常不得不局部地直接应用操作系统及其文件管理系统。

1. 数据管理技术的发展

从计算机发展史上看，数据管理技术经历了人工管理、文件系统和数据库系统三个发展阶段（表 3－2）。在早期的人工管理阶段，数据属于具体的应用程序，即使是同一种应用的用户需要很多相同数据时，也要各自具备自己的数据；由于数据相互间不能共享，不得不重复存储，导致数据大量冗余。当数据管理技术发展到文件系统阶段时，数据记录内已有结构，可通过高级语言等途径

存取，用户开始可以共享数据。例如，不同的应用者可以利用高级语言读取§2.2.3中表2-1的DLG数据文件中的每个数据。但是，这时的数据和记录没有较大数据集的整体结构，面向某种应用的烙印仍然比较明显，数据与行业应用程序之间还有依赖关系，数据的独立性较差，共享性差，冗余度大。

表3-2 数据管理技术三阶段的比较

	人工管理	文件系统	数据库系统
数据管理者	用户(程序员)	文件系统	数据库系统(DBMS)
数据面向和服务	面向某应用程序	面向某一种应用	面向现实世界有关普遍性特征
数据共享程度和冗余	无共享，大量冗余	共享性差，冗余度大	共享性高，冗余度小
数据的独立性	无，完全依赖于程序	独立性差	较高的独立性
数据的结构化	无	记录有结构，整体无结构	整体结构化，用数据模型描述

数据管理技术到数据库系统阶段才真正跃上台阶。首先，数据管理面向现实世界中各种数据组织以及数据间联系的普遍性特征，来建立数据模型，并基于数据模型将数据结构化，形成能由DBMS统一管理，面向普遍应用的由相互间有机联系的数据项、记录和文件所组成的数据库整体结构。这样，数据管理技术不再是面向某一种应用，DBMS软件可为各行各业所应用，共享性明显提高，冗余度相应减小，同时数据的独立性也明显提高，数据库中的数据与用户应用程序相互独立，应用程序不因数据性质的改变而改变，数据的性质也不因应用程序的改变而改变。

正因此，数据库里的文件管理，与文件系统的管理的内涵完全不同，数据库的文件和数据有内在联系，数据库系统透过文件，全面有效地实施数据管理。数据库中的数据可长久存储，并能为应用程序交叉使用。数据库还提供改善访问性能的高效结构和算法，包括索引管理、数据聚集、数据缓冲区、查询路径选择及查询优化等；采用高级查询语言访问和查询实体信息；通过事务管理能监视数据库的交互过程，确保数据的一致性(正确性)和稳定性；数据库还不断发展自己的分析统计功能等。总之，数据库系统使数据操作和管理的效率大大提高，数据查询迅速、准确，使数据管理进入高级阶段。

2. 数据库系统的进一步发展

数据库系统技术不仅解决了上述数据管理的基本问题，近十余年来，还在继续向更高水平快速发展。为了更高质量地进行数据服务，数据库技术，特别是大型商业数据库，形成了一系列较高级的数据管理功能。这些高级功能可归纳为三个主要方面。

• 数据保护功能。数据保护是高质量数据管理服务的重要保证，特别对银行、电讯等公众服务部门，数据保护职能具有生死攸关的意义。数据保护主要

有四个方面的内容：首先是安全性控制，即防止数据丢失、错误更新和越权使用。数据库的用户通常只能使用和更新某些数据，只有数据库管理员才能对整个数据库进行操作。其次是完整性控制，即保证数据正确、有效和相容。第三是并发控制：就是既要能做到同一时间、周期内允许对数据的多路存取，又要能防止用户之间的不正常的交互作用。第四是故障的发现和恢复，要提供一套措施，警惕和发现故障，并在发生故障时，尽快自动恢复数据库的内容和运行；当系统突遇事故而发生故障甚至毁坏时，它具有将系统恢复到事故前某种特定状态的能力。

• 网络化数据管理功能。20 世纪 80 年代，随着网络技术的迅速发展，对通过网络传输和存取数据的需求迅速增加，网络化的数据管理技术得以发展起来。首先值得一提的是基于文件方式的客户/服务器技术，该技术将文件及数据存放在服务器中为网上用户共享，网上各个客户端上的用户都可以向服务器存取文件。但这种基于文件方式的客户/服务器技术有很多不足，除文件方式本身管理数据的缺点外，它的网络运作效率也较低。80 年代后期，数据库系统形成了自己的客户/服务器技术。基于大型商业数据库的客户/服务器技术不仅具有较高级数据管理功能（包括上述高效、安全、稳定、可靠、协同、并发操作等优点），而且网络运作效率高，因为这时在网上传输的不再是一个个文件，而是服务器中经数据库系统（根据用户请求）挑取的数据。基于大型商业数据库的多级客户/服务器模式业已成为现代大型机构内网络管理的主要方式，并在 90 年代中后期被 GIS 采用。

20 世纪 90 年代以来，全球范围的无序网络——互联网（Internet）技术飞速发展。分布在世界各地的企事业单位、机构和组织的内部网络（Intranet）纷纷与互联网相连，形成"分布式"的全球网络环境。为了适应新的 Internet/Intranet（一般译为互联网/企业网）环境，数据库系统进一步开发新的网络功能，向"分布式数据库系统"发展。

• 数据仓库和数据挖掘、知识发现技术。数据仓库（data warehouse）、数据挖掘和知识发现是 20 世纪 90 年代中后期迅速发展起来的一种数据库新技术。数据挖掘和知识发现（DM & KD 技术，已见本书 §1.1.3.5）是从巨型或海量数据库中系统地上规模地提取信息、发现知识的技术；而数据仓库正是面向数据挖掘和知识发现的新的数据库组织形式。一般认为，数据仓库是面向主题的、集成的、不可更新的，随时间不断扩展的数据的集合，这种集合支持数据挖掘和知识发现，面向企业的分析决策的应用。

3. 文件管理系统

随着数据库系统的发展，文件系统原有的数据管理职能逐渐淡化，不难理解，有了统一高效管理数据的数据库，一般而言，就没有必要再采用文件方式去

管理数据。但是，文件系统本身至今仍有存在价值，因为操作系统需要文件管理功能 例如，现代 Windows(视窗)的任何用户可看到：文件管理系统是 Windows 不可缺少的重要部分，否则，Windows 将很不方便应用。当然，现代文件系统的职能与数据库技术形成前的文件系统已明显不同，现代文件系统主要管理文件；除了 GIS 等少数领域，已没有什么人用文件系统去管理数据。

现代文件管理系统指数据库以外，直接由操作系统管理的文件系统。文件是某种类型逻辑记录或相同性质记录的集合，存储在硬盘等外存设备上；记录是文件中可存取的基本数据单位。文件系统的主要职能是解决文件中记录在外存设备上如何安排和组织，以及实施对文件及其记录的访问方式和修改等问题。为了提高查找、存取和修改的效率，记录在计算机中存储的物理方式常常并不按文件中记录的逻辑顺序(用户应用的记录顺序)来安排。从这个意义上，基本的文件组织方式可以分为顺序文件、索引文件、直接文件和倒排文件。每种方式各有优缺点。

顺序文件是记录之逻辑顺序和物理(存储)顺序一致的文件，是最简单的文件组织形式。索引文件是带有索引表的文件，索引表中列出记录的关键字和记录在文件中的位置(地址)。读取记录时，只要提供记录的关键字值，系统通过查找索引表获得记录的位置。直接文件也称为随机文件或哈希文件，其存储是根据记录的关键字之值、通过某种转换方法得到一个物理存储位置，然后把记录存储在该位置上。查找时，通过同样的转换方法，可直接得到所需要的记录。倒排文件中的记录不仅有主关键字，还安排有按辅关键字或多关键字(多码)，倒排文件不按主关键字，而是按辅关键字或多码来查找。倒排文件在地理信息系统中有特殊意义，因为 GIS 实践中常需要依据一些非关键属性，如坡度、土壤等来提取数据。这时，为提高查找效率，缩短响应时间，需要仔细分析辅关键字，建立一组辅索引。

3.1.3 关系数据库系统技术

关系数据库技术是目前应用最广泛的，也是地理信息系统目前所主要应用的数据库技术。

1. 传统数据库系统技术的发展

不仅数据管理技术经历了人工管理、文件系统和数据库系统的发展阶段，数据库技术本身也基于不断演化的数据模型而发展，而经历了几个发展阶段。不同阶段数据库技术的名称，通常就以所基于的数据模型来称呼，正如数据模型通常以它所基于的数据结构名称来命名一样(见 §2.1.1.2)。研究最早的两种数据模型是层次模型和网状模型，相应的数据库技术就称为层次数据库和网状数据库技术 这两种技术现在虽然由于本身的不足而很少应用，但是对数据库技术发展做出了贡献：层次数据库是 DBMS 的先驱，网状数据库是数据库概念、方法和技术的奠基。现在一般认为，层次和网状数据库是第一代数据库系

统技术。

第二代数据库系统就是基于关系模型的关系数据库技术，该技术出现于 20 世纪 70 年代，80 年代走向成熟并广泛应用，90 年代进一步发展成为现代应用最广泛的数据库技术。目前，数据库技术正在向第三代发展。

第一代数据库常称为非关系数据库，相应地，层次模型和网状模型合称为非关系数据模型。非关系数据模型依照现实世界中的某些联系来建立记录类型间的联系。这些联系包括："一对一"、"一对多"和"多对多"三种联系。例如，学校与校长这两个记录型是一对一的联系；学校与学生是一对多的联系（相当于包含关系）；学生与课程是多对多的联系。

非关系数据库中的层次模型和网状模型的基本区别在于：层次模型基于数据结构中的"树形"结构；而网状模型将数据组织成有向图结构（请参阅任一种"数据结构"教材）。§2.3的图 2-14 中的四叉树就是一种树形结构，我们可以参照该图来理解。树形结构明显地表达一种层次关系，即一对一和一对多的联系，因为该图中每个结点（除根结点外）有且仅有一个父结点。如果将这个结点条件放宽，即结点可以有任意个父结点（一般通过指针来实现这一点），那么，图 2-14 便会变成网状交错的结构，因而能进一步表达多对多的联系。非关系数据库的 DBMS 存在着结构比较复杂操作也比较复杂的问题。

2. 关系数据库技术

关系数据库基于关系模型来组织数据。在 §2.1.1.2 中说过，数据模型通常由三要素组成：数据结构、数据操作和完整性约束条件。在关系模型中，数据结构（在用户看来）就是一张满足一定条件的二维表，一张表亦称为一个关系。表 2-3 就是一张二维表。参照该表可见，二维表中的列，称为字段、分量或数据项，各表示一种属性；表中的一行称为一个元组或记录，表达现实世界中的一个实体。例如，表 2-3 中每一记录表示一条线，各数据项表达这每条线的各有关属性。二维表数据在计算机中存储可采用顺序存储结构或链式存储结构实现，两种方式各有优缺点。

在外存储器中的实际关系数据库是很多关系表的集合，这些表称为基本表。但用户实际可应用的关系表远不至此，因为在关系数据库中，可以很方便地从基本表中导出新的关系表（参见图 3-1），只要两个或以上的表具有公共属性项，就可以通过公共属性项建立起两个关系表之间的联系，导出新表。例

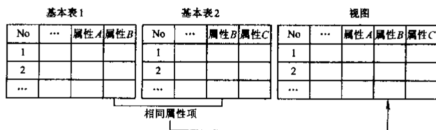


图 3-1　基本表与视图示例

如学生成绩表、学生助学金表和学生健康表中都有属性项“学生姓名”,那么,学生的成绩、助学金和健康属性就可以形成一张新表,直接相互联系和对照。又例如,区教育局数据库中,若各学校的学生健康表中都有“视力”属性,那么,就可以导出全区学生视力表,进而可得全区视力在0 5以下学生数据;若学生情况表中有教师(班主任、任课教师等)属性,还可以通过学生查找教师情况等。

导出的新表不存储到外存中,因而被称为“视图”

关系模型能用二维关系表来表现现实世界中的多种联系。在一张具体的关系表中,例如学生的综合情况表中,不仅有同一类各实体(即记录)间的联系,而且有这类实体的不同属性间的联系(如果表中有两种或以上属性的话)。不仅如此,通过视图又能给出不同表格间的不同类型的实体和属性的联系。由此可见,关系模型虽然把现实世界中的联系归结为简单的二维关系表,但它能体现现实世界中的多种联系,包括一对多和多对多联系等,尽管有时难以处理得较为直接或简练。例如,学校和学生之间的一对多联系,只需在学生表中给出“学校”的属性,并在每个学生记录中都填上该学校名,就可以实现。当然,这样做有不少数据冗余。

关系模型的数据操作可利用关系代数和关系运算来实现。20世纪80年代,关系数据库技术正式形成了一套DBMS的标准数据库语言,即SQL语言。SQL语言不仅具有丰富的数据库操作功能,而且具有数据库定义功能和数据控制功能,是集数据操纵、数据定义和数据控制为一体的关系数据语言。自从SQL成为国际标准语言之后,各个数据库厂商纷纷推出各自的SQL软件或与SQL的接口软件,SQL成为共同的数据存取语言和标准接口,使不同数据库软件之间的互操作成为可能。这是数据库技术的又一革命性飞跃,也是对计算机技术的重要贡献,对其他计算机分支技术也产生了很大影响。

关系模型的主要优点是:结构简单清晰,没有复杂的存取路径,数据修改和更新方便,容易维护和理解,易学易懂。关系模型以简捷的二维表方式表达现实世界中多种联系,并支持数据的重构(视图),其数据描述具有较强的一致性和独立性。关系模型具有较严密的数学(关系代数和关系运算)基础,具有一定的演绎功能。此外,关系模型还具有关系结构稳定,表中具体内容可不断变化的特点。

关系模型的上述优点很大程度来自于它的“规范化”要求或特点。原来、关系模型在数据库技术的第三要素(即约束条件)方面要求较高,它提出不同层次的“范式”要求。范式的定义涉及到数据库理论的深层次内容,本书不可能涉及,但又无法回避,因为关系模型的规范化要求正是导致空间数据管理的难处之一。为此,我们将关系模型的规范化要求通俗地解释为:关系数据库必须具有非常规整的结构,在一张关系表中,所有的元组或记录都是同质的,即有(类

型、长度等)相同的属性项,每一个分量都不能分割,整个表格整齐划一等(请再参照表2-3)。关系模型的规范化要求,为关系模型带来简明、方便、稳定、一致性等优点,但也导致某些理论和应用上的局限,包括GIS应用的难处。

3 关系模型的局限和不足

目前,关系模型及其应用还存在如下问题:

• 运作效率不够高。例如,重新构造数据的操作相当费时;实现关系之间联系需要执行系统开销较大的联接操作。

• 描述对象语义的能力较弱。现实世界非常复杂,数据种类和数量繁多,许多描述对象本身具有复杂的结构和涵义,为了用规范化的关系描述这些对象,需要将对象进行不自然的分解或重复存储(例如上小节提到的学生与学校关系之例),从而在存储模式、查询途径及其操作等方面均显得不甚理想。

• 控制和操纵复杂对象的能力较弱。由于复杂对象的表达只能通过关系的规范化分解来实现,过多的不自然分解必然导致控制和操纵的困难。

• 表达、管理复杂对象的能力不足。不直接支持一对多关联的层次结构;不直接支持概括、分类和聚合的模拟。常用的、基本的规范化要求(第一范式)不允许嵌套元组和嵌套关系存在。为了表现和模拟现实世界中的嵌套结构(GIS中常见,如多层次嵌套的多边形),需要扩展模型。关系模型的数据类型简单、固定,难以进而扩展或抽象出新数据类型,因而不具备管理多种类型数据对象的能力。

• 可扩充性较差。在关系模型中,导出的关系与原有的关系相互独立;规范化要求对属性和数据项提出较高的规整性要求等,也皆不利于系统的扩充。

3.1.4 面向对象数据库系统

1. 新一代数据库技术

关系数据库的巨大成功,导致应用领域迅速扩展。极其广泛的应用领域,提出了很多传统数据库技术解决不了或难以解决的课题,从而反过来推动数据库技术的进一步发展。目前一般认为,数据库技术的新发展有下述三个主要特点:

第一,面向对象方法和技术对数据库发展影响最为深远。

第二,数据库技术与多学科技术,包括计算机其他各分支技术相结合。

第三,突破传统数据库技术,面向各种应用领域的数据库技术研究。其中,GIS数据库技术研究就是这方面最引人注目的领域之一。

1990年高级DBMS委员会发表了"第三代数据库系统宣言"。经多年研究讨论,对第三代数据库系统的基本特征已有了共识。这些基本特征是:

第一,第三代数据库系统技术应当支持数据管理、对象管理和知识管理,必

须支持面向对象数据模型。

第二、第三代数据库系统必须保持或继承第二代数据库系统的技术。

第三、第三代数据库系统必须具有良好的开放性：支持数据库语言标准，支持标准网络协议；具有良好的可移植性、可连接性、可扩展性、互操作性等。

这里两次提到面向对象方法或模型，由此可见面向对象方法在新发展中的重要意义。为此，有必要在下面专门解释。但它们是介绍性的，不作基本要求。

2. 面向对象数据模型

为了较好地模拟和表达现实世界中的复杂现象，克服传统数据模型的局限性，人们从更高的层次提出了一些数据模型。其中，面向对象(object-oriented)数据模型是高层次数据模型的最重要发展，因为它包含了其他模型在数据模拟方面的很多概念，并能很好地模拟和操纵复杂对象。面向对象数据模型以面向对象概念和面向对象程序设计为基础。面向对象设计也称为面向对象方法，是为了克服软件质量和软件生产率低下而发展起来的一种程序设计方法，目前应用非常广泛。面向对象方法的基本思想是：对问题领域进行自然的分割，以更接近人类通常思维的方式建立问题领域的模型，以便对客观的信息实体进行结构模拟和行为模拟，从而使设计出的系统尽可能直接地表现问题求解的过程。面向对象的定义是指无论怎样复杂的事例都可以准确地由一个对象表示，这个对象是一个包含了数据集和操作集的实体。

面向对象数据模型具有下述主要特征：

• 对象与封装性。在面向对象的系统中，所有的概念实体都可以模型化为对象。多边形地图上的一个结点或一条弧段是对象，一条河流或一个省也是一个对象。一个对象由描述该对象状态的一组数据和表达它的行为的一组操作(方法)组成。例如，河流的坐标数据描述了它的位置和形状，而河流的变迁移动表达了它的行为。由此可见，对象是数据和行为的统一体。

面向对象模型支持对象标识，这意味着它提供诸如对象赋值、对象拷贝以及对对象标识和对象等进行的测试操作。传统关系模型是基于值的系统，它不支持对象标识的概念。然而一些扩展的关系数据库或许多地理信息系统在基于值的模型中通过引入显式定义的对象标识符来模拟对象标识。此时，对象标识号实际上是作为一个属性值存于关系表中。

• 分类。类是关于同类对象的集合。具有相同属性和操作的对象组合在一起形成类(class)。类描述对象的属性特征以及作用于类中对象上的操作方法。每个对象都是这个类的一个实例，它们既有共性，也有个性，前者是它们成其为一类的基点，属于同一类的所有对象在内部状态的表现形式上(即型)相同，可共享相同的属性和方法，但它们可以有不同的内部状态，即有不同的属性值。类中的对象并不是一模一样的，而应用于类中所有对象的操作却是相同的。因此，在实际的系统中，仅需对每个类型定义一组操作，供该类中的每个对象应用。但因每个对象的内部状态不完全相同，所以要分别存储描述每个对象个性的属性值。

以一个城市的 GIS 为例，它包括了建筑物、街道、公园和电力设施等类型。建筑物类中，每个对象都有建筑物用途、地址、房主、建筑日期等属性，并可能需要显示建筑物、更新属性数据等操作。但每个具体的建筑物有各自的属性值，如解放路 3 号楼，房主为李汉，建筑日

期为 1975 年等。每个建筑物都使用建筑物类中操作过程的程序代码，代入各自的属性值操作该对象。

• 概括。在定义类型时，将几种类型，例如 class 1 和 class 2 等中某些具有公共特征的属性和操作抽象出来，形成一种更一般的超类（superclass）。这就是概括。class 1 和 class 2 等称为超类的子类（subclass）。

子类与超类的关系是 is a 的关系。例如建筑物是饭店的超类，因为饭店也是建筑物子类还可以进一步分类，如饭店类可以进一步分为餐馆、普通旅社、涉外宾馆、招待所等类型。所以一个类可能是某个或某几个超类的子类，同时又可能是几个子类的超类。

建立超类实际上是一种概括，避免了说明和存储上的大量冗余。但是它还不足以描述子类、超类间的状态和操作联系。为此，还需要一种机制，使子类对象自动获得它的超类的状态和操作。这就是面向对象方法中著名的模型工具——继承。

• 继承。继承是一种服务于概括的工具，它可将超类的某些属性和操作遗传给子类例如，在前面概括的例子中，饭店类是建筑物类的子类；建筑物类的一些操作，如显示和删除对象等，和建筑物类的一些属性，如房主、地址和建筑日期等，可以仅在建筑物类中定义，然后遗传给饭店类等子类。在遗传的过程中，还可以将类的属性和操作遗传给子类的子类。例如可将建筑物类的一些操作和属性通过饭店类传给孙子类（如招待所类）等。继承是一有力的建模工具，它有助于进行共享和应用，提供了对世界简明精确的描述。

继承有单个继承和多个继承。单个继承是指子类仅有一个直接的父类，而多个继承允许多于一个的直接父类。GIS 中经常要遇到多个继承的问题。例如，运河或河流既可以是交通线类的子类，又是水系类的子类，因而可以继承交通线类和水系类两方面的特性；同样，饭店既可以继承建筑物类的特性，又可以继承旅游设施类的特点。

• 关联或联合。关联或联合是将同一类对象中的具有部分相同属性值的对象组合起来，设立一个更高水平的对象表示这些有相同属性的集合。例如，在农用地类别中，具有相同权属的农用地（宗地）与相邻宅基地可以构成一个高一级的组合。这里，联合与概括的概念不同：概括是对类型进行抽象概括；而联合是对对象进行抽象联合，联合的分子对象同属于一个类型。联合所得到的对象叫组合对象，组合对象与其分子对象之间的关系是 member-of 的关系。

• 聚集与复合对象。聚集有点类似于联合，也是将几个不同特征的对象组合成一个更高水平的对象；但是，组合成的对象，称为复合对象，与其分子对象之间的关系是 parts-of 的关系。例如，房屋可以是墙、门、窗和房顶组成的复合对象；城市可以是街道、居民区、商业区、公园等的复合对象。在聚集关系中，分子对象是复合对象的一部分，它们有自己的属性描述数据和操作，这些属性和操作不能为复合对象所公用，但复合对象可以从它们那里派生得到一些信息，从分子对象中提取部分属性值。复合对象一般不继承子对象的操作。在许多文献中，联合的概念附在聚集的概念中，并将组合对象和复合对象合称为复杂对象，并且都使用传播的工具提取子对象的属性值。

• 传播。传播是服务于联合和聚集的工具，它通过"函数"般的功能，将子对象的属性信息传播给复杂对象。就是说，复杂对象的某些属性值不单独存于数据库中，而是从它的子对象中提取或派生。例如，多边形的位置坐标数据，并不直接存于多边形文件中，而是存于

弧段和结点的文件中,多边形文件仅提供一种组合对象的功能和机制,通过"传播"提取到多边形的位置信息。通过传播,还可以在联合和聚集中获取派生的信息。例如,一个县的人口可由关联的乡镇人口求和得到;一家旅社的客房总数和职工总数可以从客房数据和职员数据中派生得到。这样做还可以保证数据库的精练性和一致性,因为独立的属性值只需存储一次,不会因操作或更新而导致不一致性。

综上所述,面向对象数据模型描述和管理客观世界复杂实体及其相互联系的能力,比关系模型等传统数据模型要强得多;又由于它是在面向对象程序设计方法基础上发展起来的,在实现计算机运作方面也具有很多传统数据模型所没有的优点。

3. 面向对象数据库系统

面向对象数据库系统就是采用面向对象方法建立的数据库系统,它首先具备常规数据库系统所提供的全部特性,其中包括:数据模型、永久性、并发性、事务管理、恢复功能、查询语言、辅存管理、安全性等。除此而外,根据上述面向对象模型的特点,不难想象,面向对象数据库系统还应有下列优点:

- 数据抽象。它能开发和使用灵活的,逻辑上复杂的抽象数据类型。
- 较强的信息模拟能力。现实世界中实体的结构和行为在计算机环境中被描述为类和对象,这种形式特别适合于模拟和用途。
- 对象标识。系统为对象定义并维护唯一的标识符,这一标识符独立于对象本身的内容和地址,并支持对象的共享。
- 封装性与数据隐藏。对象由其类型定义的方法操纵,其内部结构对于外部访问来说是隐藏的。
- 主动(智能)数据。过程与数据的封装使系统具有推理能力,并具有一致性和有效性,这也就增强了数据库定义智能操作触发器的能力。
- 继承性。由类的抽象可构成类层次,类层次中子类可继承其超类的数据与行为。超类的特殊化为子类,子类的概括化形成超类。继承性自动维护类层次中的关系。
- 多态性。数据和成员函数都可以在运行时刻动态地解释为不同的含义,这是用重载和动态联编实现的。
- 复合性。一组对象可以是另一个对象的一部分。
- 消息传递。对象之间的交互是通过消息激活对象中的方法来体现的。这个特性增强了系统的模拟能力。
- 计算完整性。系统在计算上应该是完整的。
- 可扩充性。系统提供的操作、结构和约束不是固定的,用户可以在应用中增量地定义新的操作和类型。

4. 面向对象数据库系统的实现方式

基于面向对象数据模型,面向对象数据库系统目前主要通过以下三种方式得以实现:

第一种方式是扩充面向对象程序设计语言(object-oriented programming language,简称OOPL),在OOPL中增加DBMS的特性。例如,ONTOS就是通过扩充C++而形成的一种OODBMS,它用标准C++代码定义类和函数,并提供主动数据字典的概念,使数据能动态定义。在OODBMS中增加处理和管理地理信息数据的功能,则可形成地理信息数据库系统。

在这种系统中，对象标识符为指向各种对象的指针；地理信息对象的查询通过指针依次进行（巡航查询）；这类系统具有计算完整性。这种实现途径的优点是，能充分利用 OOPL 强大的功能，相对地减少开发工作量；同时，容易结合现有的 C ++ 语言应用软件，使系统的应用范围更广。这种途径的缺点是没有充分利用现有的 DBMS 所具有的功能。

第二种方式是扩充 RDBMS（关系数据库管理系统，即 relational DBMS），在 RDBMS 中增加面向对象的特性。扩充 RDBMS 也用 OOPL（如 C ++），也可用常规程序设计语言（如 C 等）。IRIS 就是用 C 语言和 LISP 语言扩展 RDBMS 所形成的一种 OODBMS。扩充 RDBMS 途径的优点是：能充分利用 RDBMS 的功能，可使用或扩展 SQL 查询语言；能结合二者的特性，大大减少开发的工作量。缺点是数据库 I/O 检查比较费时，需要完成一些附加操作，所以查询效率比纯 OODBMS 低。

第三种方式是建立全新的支持面向对象数据模型的 OODBMS。这种实现途径并不以 OOPL 作为基础，而是从重视计算完整性的立场出发，以记述消息的语言作为基础，创建全新的数据库程序设计语言（DBPL）或永久性程序设计语言（PPL）。此外，它还提供非过程型的查询语言。02 系统就是用这种途径实现的。它由三个层次组成，模式管理（SM）、对象管理（OM）和 Wisconsin 存储系统（WISS）。SM 负责类别、消息和公共区名字的生成、查询、更新和删除；OM 负责复合对象及复合值与消息的交换；WISS 则提供构造记录的各种文档的存储方法。这种实现途径的优点是：用常规语言开发的纯 OODBMS 全面支持面向对象数据模型，可扩充性较强，操作效率较高；重视计算完整性和非过程查询。缺点是数据库结构复杂，并且开发工作量很大。

上述三种开发途径各有利弊，侧重面也各有不同。

§3.2 地理空间数据管理

本节专门讨论地理空间数据管理的几种模式以及它们的实现方法。地理空间数据管理的技术基于通用数据库的理论或理念，并直接应用后者的技术；但由于空间数据的特殊性，通用的数据模型和数据库管理系统并不能完全适应于地理空间数据的管理，GIS 还必须发展自己独特的技术和方法，以解决地理空间数据管理的特殊问题，并从特定的角度对通用数据库系统的技术和理论作出贡献。

3.2.1 地理空间数据管理不同于常规数据管理的特点

地理空间数据管理有一系列不同于常规数据管理的特点。本小节将分析这些特点，从而揭示，通用数据库系统技术之所以不能完全适应地理空间数据管理的原因。

1. 栅格地理空间数据的管理

数据管理或数据库技术通常构架在数据结构之上。地理空间数据管理的

方式也会因空间数据结构的不同而不同。如上章所述，栅格GIS和矢量GIS作为两类实用化的GIS系统，分别建筑在栅格和矢量空间数据结构上，它们都需要管理地理空间数据，建设地理空间数据库。对这两种空间数据的管理，通用数据库系统技术都不能很好适应，但由于栅格与矢量两种空间数据组织的方式不同，它们与通用数据库系统的关系还有所不同。

在§1.2.2中说过，空间数据可以看作由三个部分组成：

地理空间数据＝空间位置数据＋属性数据＋两种数据间落实于每个地物的挂联关系

在栅格数据情况下，栅格图像本身包含着组成矩阵形式的属性数据，也包含着隐含在格网的行、列数之中的空间位置（特征）数据，因而也隐含着与生俱来的、属性数据与空间位置数据之间的挂联关系。栅格数据的管理归结为图像数据的管理。地学图像处理系统多年来一般采用文件方式管理图像数据；但近年来采用大型关系数据库管理图像数据的问题已日益受到关注，其主要目的是享受大型关系数据库技术进步所带来的高效、安全、稳定、可靠、协同和并发操作等利益，实现各种海量地理空间数据的后台统一管理。当然，目前关系数据库技术管理栅格图像数据还有局限性，在图像处理操作效率上可能不及直接采用操作系统下的二进制图像文件（矩阵数组结构）。

2. 矢量地理空间数据的管理

与位置数据和属性数据天然一体的栅格数据不同，矢量数据结构的空间位置数据和属性数据通常是分立的两套数据，两种数据间落实于每个地物的挂联关系通过每一点、线或多边形的唯一标识符来实现（参见§2.2.1及图2-4）。其中，属性数据非常适合于用关系数据库技术来管理，因为属性数据与常规计算机数据类似，而且很多GIS系统中的属性数据采用所属领域规范的分类编码，国家、行业或地区的标准分类编码系统对每一种地物的类型、面积或长度等都有明确的规定。因此，一般地理信息系统都直接采用关系数据库技术来管理地理属性数据。

然而对空间位置数据的管理，通用的关系数据库技术就显得不足了。空间位置数据有两层含义：其一是地物本身的地理位置，常用某一种地理坐标系中的坐标及其组合来表达；其二是多个地物位置之间的相互关系，或空间关系，如地物之间的距离、相邻、相连和包含关系等。传统数据库技术对无论哪一方面都存在问题。

首先，关系数据库管理系统对数据记录有规范化要求。例如（第一范式）起码要求每一条记录是定长的，数据项表达的只能是原子数据，不允许分裂或嵌套记录。而空间几何（图形）数据一般不能满足这种结构化要求，用一条记录表达一个空间对象，其数据项就可能是变长的。例如，一条弧段的坐标，其长度是

不可限定的，它可能是两对坐标，也可能是 10 万对坐标。另外，空间对象可能包含或嵌套一个或多个其他空间对象。例如，一个多边形可能含有多条弧段，还可能多重地包含较小多边形等。关系数据库技术没有直接嵌套管理数据的功能，管理包含关系就相当麻烦。空间图形数据的非规范化特征是它难以直接采用通用的关系数据库系统管理的最主要原因之一。

除非规范化特征外，空间位置数据还有其他特征也限制了关系数据库的应用。例如，地理空间数据管理要考虑空间分布特征，以致除了通用性数据库管理系统或文件系统关键字的索引和辅关键字索引以外，一般还需要建立空间索引。地理空间对象之间还存在着复杂的空间关系，并且还能产生新的关系，传统数据库技术对这些关系缺乏描述能力。拓扑特性虽然能通过属性表部分地描述，从而方便了空间数据的查询和空间分析，但是，涉及拓扑关系的空间数据的一致性和完整性维护却是较复杂的。例如，拓扑结构的多边形不直接包括坐标数据，仅含组成多边形的弧段标识。因此，当涉及多边形查找、显示和分析操作时，都要操纵和检索多个（弧段）数据文件方能得以实现。

3．空间数据管理具有明显的海量、区域性和多层次特点

在 §1.2.3.3 和 4 中，我们曾较详细分析过 GIS 的海量、区域性和多层次特点，这些特点也给空间数据管理带来某些难题。GIS 数据库，即使是一个小区域的数据库，也往往是海量或超海量的。一个城市地理信息系统的数据量可能达几十个 GB（1GB = 10^9 字节），如果加上应用影像数据，则可能达几百个 GB。这样的数据量在其他一般数据库中是很少见的。地理信息系统的海量、超海量数据，带来了系统运转、数据组织与储存、网络传输等一系列技术困难，自然也给数据管理增加了难度。

地理信息系统的区域性、多层次特点，也对 GIS 数据库技术提出特殊的要求（回顾图 1－3）。首先，很多 GIS 系统都涉及多幅水平相接的地图，这就要求 GIS 数据管理技术具有图幅管理的功能，方便用户进行跨图幅的运作。其次，GIS 系统一般都要涉及两个或以上区域层次的管理，因而又要求 GIS 具有不同比例尺地图的数据管理功能。再者，GIS 系统通常还要涉及多种专题图层的管理，这就要求 GIS 还必须具备专题图层管理的功能。这样，GIS 数据库必须解决大型图库管理的问题：在横向管理好图幅，在纵向管理好区域的或专题的图层。

4．空间数据管理涉及到较复杂的实体类型和空间操作

地理空间数据管理还面临一系列至今未能解决好的问题。首先，地理空间数据的实体类型繁多，不少对象相当复杂，地理空间数据管理技术还必须具备对地理对象（大多为具有复杂结构和内涵的复杂对象）进行模拟和推理的功能。但是，传统数据库系统的数据模拟主要针对简单对象，管理的实体类型较少，因而无法有效地支持以复杂对象为主体的 GIS 领域。随着 GIS 技术向三维甚至

更高维方向发展，GIS 系统需要描述表达的对象愈来愈复杂，这个问题将愈来愈突出。

从数据操作的角度，地理空间数据管理中需要进行大量的空间数据操作和查询，如矢量地图的剪切、叠加和缓冲区等空间操作、影像特征提取、影像分割、影像代数运算、拓扑和相似性查询等；而传统数据库系统擅长于操纵和查询文字和数字信息，难以适应空间操作。

3.2.2 传统的地理空间数据库系统

由于上述地理空间数据管理的特点，地理信息科学在应用常规数据库系统技术和理念的基础上，发展了自己独特的地理空间数据管理技术。本节讲述目前应用最广泛的传统的地理空间数据库系统技术。但在此之前，先讨论“地理空间数据库系统”的基本概念。

1. 地理空间数据库系统

当前在 GIS 领域，有关地理空间数据管理技术和数据库的用词常常比较随意。地理信息系统的数据库，有时也称为 GIS 数据库、地理空间数据库或空间数据库，这些词，加上或者不加上“技术”两字，通常就用来指 GIS 的数据管理技术，而且对其中内涵的理解见仁见智。因此，有必要对地理空间数据库系统的概念稍加讨论。我们认为：

第一，空间数据库和 GIS 数据库(技术)等简化用词，是可以用的，但应当看到，它们的含义并不一定等于地理空间数据管理技术。这是因为，GIS 领域的人谈及 GIS 或空间数据库时，常常是就其狭义而言的，指的是矢量 GIS 的数据库系统；但是，地理空间数据管理技术，从广义上说，应当包括矢量 GIS 和栅格 GIS 二者的地理空间数据管理技术，栅格型的地理空间数据也存在管理技术问题。

第二，GIS 数据库或地理空间数据库之所以常指矢量 GIS 的数据库，其原因是：矢量 GIS 由于与传统地图有直接、紧密的联系，一直是 GIS 的主流技术(参见 §2.4.1.3)。更重要的是，矢量 GIS 的数据管理在采用通用数据库技术或理念的基础上，已形成了体系，具有相对丰富的内容；而栅格 GIS 系统的数据库主要内容是采用文件方式来管理图像数据，总的说来，没有像矢量 GIS 那样形成有较丰富内容的数据管理技术体系。

第三，总的说来，按通用数据库的技术要求衡量，地理空间数据库管理技术尚没有真正达到“数据库技术”的层次，因为不仅栅格 GIS 系统，目前大多数矢量 GIS 系统也部分采取文件方式管理，而文件系统在数据管理的技术层次上还不能与数据库系统相比(参见表3-2)。当然，这绝不意味着 GIS 数据库技术不复杂，恰恰相反，正是由于空间数据的复杂性，一般的地理空间数据管理技术还达不到通用数据库在冗余度、数据独立性、共享性和易扩展性等方面的要求。

20 世纪 90 年代中后期,矢量 GIS 系统开始利用现代大型商业关系数据库来统一管理图形和属性数据(见本节后文),使矢量型空间数据管理可能达到现代关系数据库技术的层次。但是,目前采用这种新技术的 GIS 数据库还不普遍;而且问题解决得也不够彻底。

第四,既然很多 GIS(包括栅格 GIS 和矢量 GIS)系统中的数据库没有达到严格意义上的通用数据库技术要求,称之为"数据库"是否合适呢?我们的回答是肯定的。但是,地理空间数据库不宜照搬 §3.1.1.2 中萨师煊定义。我们认为,采用如下的定义比较适合:地理空间数据库是关于某区域的,以特定结构组织和存储以便有效管理的、相互关联的地理空间数据及数据文件的集合。这样的定义也能适合栅格 GIS 系统中的空间数据库。

2. 传统的地理空间数据库系统技术及其发展

GIS 数据库技术经历了多年发展和演变,大体经历了这样几个发展阶段。

20 世纪 80 年代以前。GIS 数据库技术处于初期阶段,技术上不成熟。

20 世纪 80 年代,在关系数据库技术走向成熟,应用迅速扩展的形势下,以美国 ESRI 公司的 Arc/Info 为代表的矢量 GIS 技术,基于关系数据库技术理念,并部分直接地采用关系数据库技术,提出地理空间数据管理的数据模型——地学关系模型(geo-relational data model);进而成功地开发出基于这种模型的矢量 GIS 数据库系统,并取得应用上的成功。该技术将关系数据库与图形文件混合管理。这是地理空间数据库系统技术发展史上第一次革命性飞跃,矢量 GIS 数据库技术由此开始自成体系。从那时起,不同的(矢量)GIS 厂家开发了多种具体的 GIS 数据库系统,也取得一定的进步,但是,基本上都没有突破地学关系模型的技术范畴。基于地学关系模型的地理空间数据库系统至今仍然应用于大多数 GIS 系统中。

20 世纪 90 年代中后期,GIS 数据库技术又有了一次飞跃。1996 年,美国 ESRI 公司与主流数据库技术的领头公司 Oracle 合作,开发出空间数据库引擎,该技术仍然基于关系数据库系统,但是,它突破了传统的地学关系模型,采用基于大型关系数据库的客户/服务器的网络模式,实现了图形数据和属性数据在大型商业关系数据库的后台统一管理,从而能享用大型关系数据库技术飞速进步所带来的利益(见 §7.5.2 和本节后面的 §3.2.4.1)。

目前 GIS 数据库技术正向更新一代转变。

正是由于新一代 GIS 数据库技术已经出现,上述基于地学关系模型的矢量型 GIS 数据库系统技术,即关系数据库与图形文件混合管理系统,现在常被称为"传统的地理空间数据库系统"。大家知道,在一些较古老的学科,"传统"意味着很多年以前的事情;但在计算机和 GIS 等发展特别快的新兴学科,只要新技术已出现,十几年前,甚至几年前发展起来的技术,就可能被称为是"传统技

术",尽管它还仍然在广泛地应用着。

下面将较详细讲解基于地学关系模型的传统矢量型 GIS 数据库系统技术。该技术不仅至今仍然广泛应用,而且也是进一步理解新的 GIS 数据库技术的基础。

3. 传统 GIS 数据库系统技术:关系数据库与文件系统的混合管理

基于地学关系模型的传统 GIS 数据库系统技术,采用关系数据库与文件系统混合管理的模式。该模式的要点是(参照图 3-2):

第一,点、线和多边形的空间位置(图形)数据与属性数据两者分别组织、管理与检索。因此,矢量 GIS 数据库有"属性数据库"和"图形(文件)库"之提法。后者有时也被称为空间数据库,虽然这种提法不准确(参见 §1.2.2.4)。

第二,两种数据间落实于每个地物的挂联关系通过唯一标识符或者内部连接码进行链接。

第三,用关系数据库的 DBMS 直接管理属性数据或属性数据库。

第四,借用关系模型的理念,利用文件系统,并通过高级语言编程来管理空间位置(图形)数据。GIS 系统采用高级语言编程,主要任务一是开发图形处理功能,二是利用文件系统组织图形数据和管理图形库。

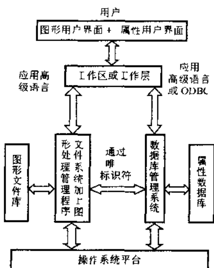


图 3-2 地学关系模型框架

第五,由于地理空间数据的海量、区域性和多层次特点,地理空间数据通常在水平方向分图幅或区域,在垂直方向上分图层来组织管理。图形文件库中的空间图形数据和属性库中的属性数据,都纳入分幅(分区)、分层组织管理的轨道。为了相应的图形数据和属性数据能按用户对具体地区和具体专题内容进行操作,需要有效地协调运作,GIS 系统还有一个非常重要的数据管理措施,这就是图 3-2 中上部的工作区或工作层。

工作区或工作层是一个或若干个图幅或图层的小型集合,其中含有组成它的图幅或图层的图形文件和属性文件(或者这些文件的信息包括目录路径等)。工作区或工作层通常与具体的工作任务相联系,用户在实际工作中打开某个工作区或工作层的过程,就是分别通过数据库管理系统和图库管理系统调出相应的属性和图形数据或文件过程。例如,在省级小流域规划管理系统中,每个小流域可以组织为一个工作区,内含组成该区域地形图、土地利用图、水利工程图

等的一系列图形文件和属性文件的目录路径;在土地资源管理系统中,工作区可按乡镇来组织等。

分幅(分区)、分层管理及相应的工作区和工作层是GIS数据管理的一个特点,不仅上述混合管理模式,现代新的空间数据管理模式也少不了它。为此,下面还将进一步讨论之。

在关系数据库的操纵和用户界面的形式上,早期的GIS与后来的GIS系统还有所不同。图3-2画出的实际上是现在的模式。在地学关系模型实践的早期,图形处理的用户界面和操纵属性的用户界面是分开的。这是因为,在图形方面,GIS系统利用高级语言编程,其内容包含制作图形用户界面和直接操纵图形数据文件。因此,图形用户界面与图形文件处理是一体的,中间没有裂缝。但是,那时的关系数据库系统不提供编程的高级语言的接口,属性用户界面和属性数据运作只能通过数据库操纵语言来实现。因此,早期的矢量GIS系统的图形用户界面和属性用户界面不得不分离。它们通过内部码连接。这样的系统通常要同时启动两个系统(GIS图形系统和关系数据库管理系统),两个系统来回切换,使用起来很不方便。

20世纪90年代,越来越多的数据库管理系统提供了C或C++语言等高级编程语言接口,使地理信息系统可以在高级语言的环境下,直接操纵属性数据,通过对话框和列表框显示属性数据,或通过对话框输入SQL语句,并将该语句通过该高级语言与数据库的接口,查询属性数据库,并在GIS的用户界面下,显示查询结果。这种工作模式,并不需要启动一个完整的数据库管理系统,实现了在一个统一界面之下的图形数据和属性数据的查询与维护。

随着数据库技术的进一步发展,关系数据库又推出了开放性数据库连接协议(简写为ODBC),使传统GIS的关系数据库和文件系统混合管理模式更加浑然一体。在ODBC推出之前,每个数据库厂商提供自己与高级语言的接口程序,以致GIS软件商不得不针对每个数据库开发与GIS的接口程序。在推出了ODBC之后,GIS软件商只要开发GIS与ODBC的接口软件,就可以与任何一个支持ODBC协议的关系数据库管理系统连接。无论是通过C还是ODBC与关系数据库连接,GIS用户都是在一个界面下处理图形和属性数据,如图3-2所示。

3.2.3 地理空间数据分区域、分层次的组织管理

1. 地理空间数据分幅(分区)、分层的组织管理

地理空间数据分幅(分区)、分层进行组织管理,是地理空间数据的海量、区域性和多层次等特点所决定的(回顾§1.2.3.3和图1-3)、GIS数据管理技术的重要特色。在计算机出现前,传统地图学的图库或地图图书馆(map library)就按“图幅”和“图层”来组织管理地图。图幅对应一块区域,一般有两种形式:标准分幅的图幅和按区域划分的图幅,例如一个县的土地利用现状图就有乡镇分幅的和1:1万(地形图)标准分幅的两种形式。在同一区域下,又有不同比例尺层次的地图,还有各种不同的专题图层。

地理信息系统中的分幅、分层管理，特别是关系数据库与文件系统混合模式的分幅、分层管理，与地图图书馆比较类似。在多数情况下，人们习惯于按不同比例尺、分幅（标准分幅或区域分幅等）、分层（专题层等）来组织图形文件及其属性数据。当然，这种管理与地图图书馆有一个很大的不同：地图图书馆中的（纸张等硬拷贝）地图是图形和属性数据的统一载体，而GIS数据库中的空间图形数据和属性数据，通常是分别存放和组织的。

20世纪90年代以来，基于面向对象空间数据模型的GIS数据库一般不分幅存贮地图，而是基于空间图形的“类”组织“无缝、无边界”的地图。在这种方式下，分幅管理表现为无缝大图上的分幅或分区索引，以满足用户对具体的局部区域和专题层进行操作、检索的需要。

正是由于GIS用户总是要对具体的区域、具体的专题内容进行操作的，无论是采用传统的关系数据库与文件系统混合模式，还是新的管理模式，GIS数据管理都需要组织工作层和工作区。

2．工作层和工作区

工作层（layer）和工作区（workspace）是面向用户具体工作服务的数据文件组织或数据处理单元。矢量型GIS系统软件通常都有这样的单元，但不同GIS软件的工作层或工作区的含义、范围和名称不尽相同。一般说来，工作区涵盖较广，一个工作区可由一系列工作层（或者它们的信息，如目录路径等）组成。当接到用户打开工作区的指令时，GIS数据管理系统就根据该工作区的组成，调出属于它的各工作层。

工作层在某些软件中也称为图层。由于不同区域层次，即不同比例尺的地图通常很少合起来运作，工作层的图层通常是一些同区域、同比例尺的专题图层，例如像图2-12左部的地形、土壤和森林那样的专题图层。

在多数情况下，工作层是真正意义上的数据处理单元。在一些GIS软件中，空间几何数据和属性数据之间的唯一标识符就是基于工作层来组织的（即每个工作层的标识符各自排序）。这样，同一工作层内的图形数据与属性数据容易相互链接，从而使工作层成为便于数据处理的单元。工作层通常由属于它的一系列图形文件和属性文件（或者它们的信息，如目录路径等）组成。当接到用户打开工作层的指令时，GIS数据管理系统就根据该工作区的文件组成，一方面索取所需的属性数据及文件，另方面获取相应的图形数据或文件，并将二者组合起来形成所需的图幅或图层，供用户处理、使用。

由于工作层和工作区是为具体任务服务的，它们是用户在实际工作中最经常涉及的地理空间数据（文件）组织单元，因而也是地理空间数据管理中的重要的基本管理层次。在§3.1.1.1讲过，数据管理的层次主要有数据项、记录、文件和数据库4级，但事实上，当数据库较大时，仅这4个层次是不够用的，即使

在通用数据库软件中，文件和数据库之间也通常增设管理层次，如"项目"级。地理空间数据库通常是海量的，就更需要文件和数据库之间的中间层次，工作层和工作区就是"文件层次"之上的数据管理层次（当然它们也身兼数据处理或操作单元的角色）。有的GIS软件还在工作区之上进一步增设更大范围的"项目"级数据文件管理层次。总之，数据库愈大，一般就需要愈多的管理层次，没有这些层次，海量地理数据库的管理将难以想象。

3. 主要GIS软件的工作层和工作区

不同GIS软件的工作层和工作区含义或范围不尽相同。一些软件的工作区的范围为一个图幅，另一些可含多个图幅。后者如ESRI公司的Arc/Info软件。Arc/Info的工作区（一个目录）结构比较灵活松散，可含多个图幅或工作层。Arc/Info的工作层称为coverage，是一种比较典型的工作层。一个coverage常表现为一个目录，目录内包含有控制信息文件，标识点文件，以及点、线或多边形的图形文件和属性文件等。但某些版本（如现在的Arc/Info 8.0、8.1版）的coverage的某些信息另放在工作区下的一个"Info"目录中。coverage大体相当于专题地图，但有时比专题地图含义狭窄。后者是因为Arc/Info中，"点"图元（point feature）既用以表示点状地物，又用以表示面状地物的标签点（label）。因此，一个coverage中，不能同时有点状地物的属性表和面状地物的属性表。这样，Arc/Info中的工作层coverage，不可能像专题地图那样，同时表达点状地物和面状地物。

ESRI公司的另一个主要GIS软件（或桌面绘图软件）ArcView和Intergraph公司的GeoMedia的工作区，结构也比较灵活松散，可含多个图幅及其工作层但ArcView的组织层次多一层。ArcView的工作区，称为view，可含若干个shape图层（或coverage层）；shape图层是ArcView的主要工作层，由几个（一般5个）图形和属性文件组成。在View层次之上，ArcView还构架了一个project（项目）层次，一个project可由若干个view组成。project显然是更高层次、更大范围的工作区。

一些从CAD软件发展或脱胎而来的GIS软件将工作区中的所有地物，无论点、线和面状地物，皆组合在一个图形文件中。换言之，与上述Arc/Info等不同，这一类软件的工作区是一个大型图形文件，只含一幅地图，不能灵活松散的组合，而且同它所含的工作层或图层一样，它本身也具有鲜明的数据处理或操作单元的功能。例如，Autodesk公司的AutoCAD Map软件和ESRI公司的ArcCAD软件中的DWG文件，以及Intergraph公司的MGE软件中的DGN文件等，就是这样的图形文件（但MGE的工作层的定义相对比较灵活）。不难想象，这些图形文件有时是非常大的，其下可包含很多的工作层或图层。

另一种国内广泛应用的GIS软件（或桌面绘图软件）MapInfo，与上述第二

种类型比较类似。MapInfo 的工作区(MapInfo 中文版称之为工作空间)一般也只对应一幅地图,它也可以包含很多图层或工作层,后者称为"工作表"。但是,MapInfo 的工作区或工作空间不是一个具有数据处理或操作单元功能的图形文件,它只是一个较简单的,记录所含工作表信息、路径等的文件;当用户指令打开某工作空间时,MapInfo 根据这些信息调出它所包含的一系列工作表,即图层,供用户使用。工作表是 MapInfo 的主要工作单元。

近年来的国产 GIS 软件也很有创意。例如武汉大学(原武汉测绘科技大学)开发的国产 GIS 软件 GeoStar 中,一个工作层就是一个 GDA 文件,GDA 文件可以是某一层地物,例如道路网,甚至是某一类地物,也可以是多层地物,甚至可以包含一个工作区的所有各层地物。GeoStar 不像 Arc/Info 的 Coverage 那样有较多的限定,这有利于提高作业效率方便工作。如果一个工作层包含的内容很多,为了显示、制图和查询方便,GeoStar 还可以根据用户需要任意定义逻辑层。一个逻辑层可以包含多个图层或地物类,而且允许交叉,例如河流可以被包含在水系层中,也可以包含在交通层中。在这里空间数据的物理存储关系没有改变,仅是建立了一个对照表,每个逻辑层包含了那些指向地物类的指针。如此等等。

4. 图库管理和空间索引

有了工作区或工作层这样的地理空间数据管理层次和数据处理单元,GIS 数据管理的工作仍然没有完结,因为在大型数据库管理中,还需要建立"图库管理"的职能。较小型的 GIS 系统及其数据库不一定需要图库管理。在一些图幅数量不大的 GIS 项目中,例如,在生物多样性保护研究中,当栖息地范围不大,比例尺要求不高时,利用一个或几个工作区或工作层进行管理就够了。但是,当地理信息系统所管理的区域和所要求的比例尺都较大时,例如在城市规划管理信息系统和土地产权产籍(地籍)信息系统中,GIS 数据库会包含大量的图幅,因而会涉及很多个工作区、工作层的数据组织和管理。有时一个 GIS 系统会包含几百、几千,甚至上万个工作区及很多工作层。这时,GIS 软件必须让用户能在整个区域内进行众多图幅(分区)、图层的调用,任意图幅的拼接和跨图幅的剪切、开窗、放大,以及跨图幅图层的漫游、查询、分析和制图等。这就是图库管理的职能。在无缝大地图的方式下,图库管理职能通过有效的分幅(分区)、分层的空间索引,以满足用户对具体的局部区域和专题层进行操作、检索的需要。

图库管理是海量空间数据管理的需要,是大型 GIS 软件的必备功能。海量空间数据管理的效率是衡量 GIS 软件优劣的重要指标之一。Arc/Info 是有效进行海量数据管理的较优秀软件之一。Arc/Info 在 8.0 版以前采用传统方式进行图库管理,其模块称为"地图图书馆"(map library)和 ArcStorm。Arc/Info 8.0 采

用全新的数据存储和管理模式,实现了空间图形数据与属性数据在大型关系数据库中的后台统一管理,相应地,它建立起强有力的空间索引和查询机制。

图 3-3 是一个空间索引的例子,该例在工作区涉及的范围建立二维空间的图幅索引。它通过一个记录每个工作区范围的空间索引文件,建立整个区域与工作区的关系。例如 W34 的范围坐标是(13 000,12 000,14 000,13 000)。建立了这样的工作区索引文件以后,用户可以在系统界面下,开窗任意进入某一个或某几个工作区。

4 000	W41	W42	W43	W44	W45
13 000	W31	W32	W33	W34	W35
12 000	W21	W22	W23	W24	W25
11 000	W11	W12	W13	W14	W15
10 000		11 000	12 000	13 000	14 000

图 3-3 工作区索引(引自龚健雅)

除图幅索引外,在图库管理中还可以建立其他形式的空间索引。

除空间索引外,现代海量数据管理还涉及并发控制、网络传输和数据格式交换等多种管理。

5. 属性文件的组织

属性数据虽然一般均由关系数据库管理系统管理,但是它的文件组织方式也要服从 GIS 系统的工作区、工作层和图库的管理要求,以便与图形文件协调运作,共同组成工作区、工作层和进行跨图幅操作。因不同 GIS 软件而异,属性文件的具体组织方式可分为三种类型。

第一种是 Arc/Info 类型。其属性数据文件建立在工作区(coverage)目录之下。无论一个 coverage 包含多少地物类,其目录仅含一个记录线属性的 AAT 表,和一个记录多边形(或点状地物)属性的 PAT 表(包括目录路径信息)。AAT 表和 PAT 表都可以扩展,即新的属性信息可以链接到表中。

第二种是 MGE 的属性数据管理方式。一个地物类对应于一个属性表文件,而且所有属性文件都在工程(项目)的目录下,即是说,不同工作区相同的地物类型的属性是放在一起的,这样便于属性的工程管理,在工程中查找某一属性要快速得多。值得指出的是 MGE 并不要求每个地物类都带属性表,例如陡坎等无关紧要的地物,可不带属性。

第三种是 GeoStar 的属性管理模式。它结合前两者的优点。在 GeoStar 中,可以设计成一个地物类为一个属性表,也可以是多个地物类共用一个属性表。这样会带来许多方便。例如,高速公路、一级公路和乡镇公路,它们的地物类型

编码可能不同，但它们的属性项可能相同，它们可以共用一个属性表，以便于查询、显示和进行最佳路径分析，GeoStar属性数据文件的组织与MGE基本类似，在建立工程之前，属性数据文件位于与工作区平行的目标，在建立工程之后，位于工程的目录之下。一个属性文件包含了该工程内所有同类空间对象的属性，在这种情况下，由于属性数据文件很大，建立关键字的索引是必要的。

在图库管理中，除图幅索引外，还可以建立其他形式的空间索引。图幅索引实际上是一种比较粗的空间索引，它根据鼠标在工程中的空间位置，迅速地找到鼠标所在的工作区。但是，如果工作区数据量较大，特别是在无缝的空间数据库中（不分图幅，一幅大地图覆盖整个系统），需要建立较细致的空间索引。否则，查询某一个地物时，欲判断鼠标落到哪一个空间地物上，需要对整个工作区中的空间对象进行（鼠标）点与点的距离比较，点在线状地物上的判别，以及点在多边形内的判别等一系列比较复杂和费时的计算。因此，很多GIS系统在工作区内建立空间索引，采用无缝空间数据库的GIS系统还需要在整个区域内建立空间索引，以便提高开窗图形操作和空间查询的效率。下面介绍几种常见的空间索引方法。

- 对象范围索引。这种索引在记录每个空间对象的坐标时，记录每个空间对象最大最小坐标。这样，在检索空间对象时，根据空间对象的最大最小范围，预先排除那些没有落入检索窗口内的空间对象，仅对那些最大最小范围落在检索窗口的空间对象进行进一步的判断，最后检索出那些真正落入检索窗口内的空间对象。这种方法没有建立真正的空间索引文件，而是在空间对象的数据文件中增加了最大最小范围一项，依靠空间计算来进行判别，以排除部分空间对象。可见，该方法仍然需要对整个数据文件进行检索，有些对象判别的复杂计算仍不能避免，因而空间检索所花费的时间仍然较多。不过，由于计算机的速度越来越快，这种方法一般也能满足查询检索的效率要求。
- R树空间索引。在上述对象范围索引中，每个目标都建立了一个范围，检索空间对象，仅检索范围与检索窗口有重叠的内容。R树空间索引方法是把这一概念进一步引申，它用一些矩形范围包含空间位置相近的目标，记录所包含的空间对象的指针，然后，以这些矩形作为空间索引。事实上，对象范围索引可以看成是最原始的R树索引方法。
- 格网索引。格网索引将工作区按照一定的规则划分成格网，为了便于建立空间索引的线性表，将空间格网进行编码，然后记录每个格网内所包含的空间对象。没有包含空间对象的网格，在索引表中没有编码，即没有记录。如果一个网格中含有若干地物，则记录每个对象的标识。如此可进行索引。
- 四叉树空间索引。四叉树结构在栅格数据表达和数据压缩方面有许多优点（见§2.3.3）。事实上，它在矢量数据的空间索引方面也很有用。其基本思想是将系统的空间对象覆盖的范围，进行四叉树的分割，然后用四叉树编码进行索引。有兴趣者可参阅龚健雅《地理信息系统基础》一书。

3.2.4 近年来GIS数据库技术的发展及展望

1. 空间数据库引擎

采用关系数据库与文件系统混合管理模式的传统GIS数据库系统技术，虽

然有重要意义并取得应用上的较大成功，但由于不得不部分地采取文件方式管理，总体上还达不到通用数据库技术在冗余度、数据独立性、共享性和易扩展性等方面的要求，更不具备现代大型商业关系数据库所发展出的较高级的数据管理等功能。因此，GIS 数据库系统技术一直在寻求更好的解决方案。1996 年起，美国 ESRI 公司与 Oracle 等主流数据库公司合作，开发出一种能将空间图形数据也存放到大型关系数据库中管理的产品，并借用“数据库引擎”的概念，取名为空间数据库引擎(spatial database engine)，简称 SDE。SDE 实现了图形数据和属性数据在大型商业关系数据库的后台统一管理，从而使空间数据管理能享受大型关系数据库技术进步所到来的种种利益。

ESRI 公司以外的 GIS 基础软件开发商和大型数据库开发商也相继开发出类似的、实现图形数据和属性数据统一管理的产品，但不称为空间数据库引擎 SDE。目前 GIS 界还没有对此类产品有公认的统称。本书中，我们将这一类产品统称为空间数据库引擎，因为这一词能较好地表达这一类产品的作用和性质。

2. 目前空间数据库引擎的两种主要方式

目前，空间数据库引擎主要有两种类型。一种以 ESRI 公司与各大型数据库公司联手开发的空间引擎 SDE 为代表，可称之为采用中间件方式的空间数据库引擎。这种技术不对关系数据库管理系统的本身作原则性扩展，而是开发一种“中间件”性质的软件，一方面将客户端 GIS 应用程序的要求转化为空间图形和属性数据的 SQL 语句，转发给 RDBMS，另方面将满足搜索条件的数据在服务器端缓冲存放，并转换为可读格式发回到客户端，其作用就像空间数据与大型关系数据库之间的桥梁。

另一类空间数据库引擎系由主流数据库厂商开发。这些数据库厂商凭借自己数据库核心技术的优势，在关系数据库管理系统底层进行某些改进和扩展，开发形成空间数据管理的专用模块，例如 Oracle 的 Oracle spatial、Informix 的 Informix Spatial Datablade 和 IBM DB2 的 IBM DB2 Spatial Extender 等。这些模块能直接存储和管理非结构化或非范式的空间数据。例如 Oracle spatial 定义了操纵点、线、面等多种类别的空间对象(新近的版本还允许用户自定空间对象)，以及操作这些对象的接口函数。这样，用户可以不通过中间件软件而直接实现地理空间数据在关系数据库中的存贮和管理。

这些数据库的空间数据管理模块近年来也为 GIS 厂商所利用。例如 ESRI 公司的 SDE 过去主要采用在数据表中加入图形数据项和多级索引、关联等手段，近年来，ArcSDE 增加了对数据库空间数据管理专用模块中的空间对象进行操作的功能，给出多种存储方案。

上述两种方式都定义了面向对象的空间数据模型，形成一种关系模型向面向对象方式扩展的局面。本书交稿前，作者领导的研究组对 ArcSDE 和 Oracle

spatial 进行了较详细的比较研究，发现两种方式都存在一定的不足。相比之下，ArcSDE 总体较优，不仅 GIS 功能较强，而且能适用于多个主流数据库系统；而数据库厂商的空间数据管理模块目前仅适用于自己的用户。

有关空间数据库引擎的具体技术内容，本书还要在 §7.4 中进一步介绍。

3. 面向对象空间数据库管理系统

面向对象数据模型最适应于空间数据的表达和管理(参照 §3.1.4)，它不仅支持变长记录，而且支持对象的嵌套、信息的继承与聚集。面向对象的空间数据库管理系统允许用户定义对象和对象的数据结构以及它的操作。这样，我们可以将空间对象根据 GIS 的需要，定义出合适的数据结构和一组操作。这种空间数据结构可以是不带拓扑关系的数据结构，也可以是拓扑数据结构，当采用拓扑数据结构时，往往涉及对象的嵌套、对象的连接和对象与信息聚集。当前已出现一些基于面向对象的数据库管理系统的地理信息系统，如 CDE 等。但由于面向对象数据库管理系统还不够成熟，价格又昂贵，目前在 GIS 领域还不太通用。尽管如此，它仍具有极大的潜在应用前景。

4. 其他发展趋势

计算机及相关领域技术的发展和融合，正在或将要为地理信息数据库系统带来革命性的变化。除了上述面向对象模型和数据库系统的应用外，多媒体技术的发展也拓宽了地理空间数据库系统的应用领域。现在广义的地理信息不仅包括图形、图像和属性信息，而且还包括音频、视频、动画等多媒体信息。虚拟现实(virtual reality，简称 VR，亦译为灵境)技术促进了地理空间数据库的可视化。VR 技术将地理空间数据被转换成一种虚拟环境，人们可以进入该数据环境中，寻找不同数据集之间的关系，感受数据所描述的环境。此外，在现代 Internet/Intranet 分布式环境下，地理空间数据库具有实现分布式事务处理、透明存取、跨平台应用、异构网互联、多协议自动转换等功能。总之，未来的地理空间数据库系统将是一个可表示复杂和可变对象的、面向对象的、主动的、模糊的、多媒体的集成数据库系统。

§3.3 地理空间数据的元数据

元数据是数据应用和管理的产物，总的看来，元数据主要是为有效地数据管理、访问和应用服务的，从内容划分上应主要归于数据管理范畴。

3.3.1 元数据一般概念

1. 元数据的由来

元数据，简单地说，就是关于数据的数据，或说明数据的数据。在计算机出

现以前的很长时期，就有很多这种性质的数据，如图书馆卡片的说明内容、出版图书的版权说明，以及纸质地图的地图类型和地图图例等。这是因为，人们在应用数据时，很自然地需要了解自己所应用数据的各种有关特点，即需要说明数据的数据。

在计算机出现后的一段时期，虽然元数据概念体系还没有形成现在这样的规模，但很多数据集都已带有“说明数据的数据”。这些说明数据常常放在数据集的头部、称为文件头）或尾部，或作为数据文件的附属文件（头文件或尾文件等）。头文件如美国农业部的土壤类型图数据（参见张超《地理信息系统》51 ~ 52 页），一套数据含 4 个文件；4 文件中，后两个文件才是真正的数据文件，而两个头文件是“关于数据的数据”。其中，头文件 1 说明数据组名（原图类型）、比例尺、数字化的地区、工作部门和时间等；头文件 2 给出图幅范围，包括图廓点的图面坐标和对应地面点的地理坐标。说明数据放在文件头的如 USGS 的 DLG 数据文件（参见 §2.2.3.2），该文件的前 15 行就是关于这一套 DLG 数据的数据，其说明内容与上述土壤数据的两个头文件类似。

随着信息社会与信息技术的飞速发展，各种数据、信息大量涌现，社会各行各业对详实、准确的各种数据的需求迅速增加，对说明性数据的要求日益提高。用户需要应用各种不同类型的数据，因而要求数据库的内容、格式、说明等符合一定的规范和标准，以利于数据的交换、更新、检索和数据库集成，以及数据的二次开发利用等。在这种形势下，元数据受到愈来愈高的重视，元数据概念也不断充实和完善。如电器和电子工程协会（IEEE）在 1994 年 2 月的白皮书中指出：元数据是关于数据的信息，它使得有效地管理和访问数据成为可能。

2. 元数据概念

元数据（metadata）一词来自希腊语，原意是关于数据变化的描述。元数据亦称为描述数据，台湾学者称之为诠释数据，是关于数据的数据，它说明数据内容、质量、状况和其他有关特征的背景信息。虽然不同学者对元数据的具体理解及所下的定义不尽相同（有的甚至将它等同于数据字典），但对元数据概念一般有如下共识：

元数据是关于数据的描述性数据，它应当是现代数据库的重要组成部分，其作用和目的是促进数据集，特别是数据库中数据集的高效利用。通过元数据可以检索、访问数据库，可以有效利用计算机的系统资源，可以对数据进行加工处理和二次开发等。

元数据对数据集的描述内容包括：对数据集中各数据项、数据来源、数据所有者及数据序代（数据生产历史）等的说明；对数据质量的描述，如数据精度、数据的逻辑一致性、数据完整性、分辨率、源数据的比例尺等；对数据处理信息的说明，如量纲的转换等；对数据转换方法的描述；对数据库的更新、集成方法等

的说明等。

由于元数据是关于数据的描述性数据，它应尽可能多地反映数据集自身的特征规律，以便用户对数据集的准确、高效与充分的开发与利用。不同领域的数据库，其元数据的内容会有很大差异。

3. 元数据的分类

进行元数据分类研究的目的在于充分了解和更好地使用元数据。根据不同的分类原则，元数据的分类体系和内容有很大的差异。

● 从不同性质、领域数据和数据库需要的角度，可将元数据分为三种主要类型：

一是科研型元数据。这类元数据的任务是帮助科研工作者高效获取所需数据。其主要目标是帮助用户获取各种来源的数据及其相关信息，它不仅包括诸如数据源名称、作者、主体内容等传统的图书管理式的元数据，还包括数据拓扑关系等。

二是评估型元数据，主要服务于数据利用的评价。内容包括：数据最初收集情况、收集数据所用的仪器、数据获取的方法和依据、数据处理过程和算法、数据质量控制、采样方法、数据精度、数据的可信度和数据潜在应用领域等。

三是模型元数据。此类元数据用于描述数据模型。其内容包括：模型名称、模型类型、建模过程、模型参数、边界条件、作者、引用模型描述、建模使用软件、模型输出等。

● 根据元数据在系统中所起的作用，可以将元数据分为两种：

其一，系统级别元数据，指用于实现文件系统特征或管理文件系统中数据的信息，例如访问数据的时间、数据量的大小、在存储级别中的当前位置、如何存储数据块以保证服务控制质量等。

其二，应用层元数据，指有助于用户查找、评估、访问和管理数据等的信息，例如，文本文件内容的摘要信息、图形快照、描述与其他数据文件关系的信息。

● 根据元数据的作用，元数据可分为两种类型：

一是说明元数据。这是专为用户使用数据服务的元数据。它一般用自然语言表达，如源数据覆盖的空间范围、源数据图的投影方式及比例尺的大小、数据集说明文件等，这类元数据多为描述性信息，侧重于数据库的说明。

二是控制元数据。这是用于计算机操作流程控制的元数据，由一定的关键词和特定的句法来实现。其内容包括：数据存储和检索文件、目标的检索和显示、分析查询及显示、数据转换方法、根据索引项把数据绘制成图、数据模型的建设和利用等。

4. 元数据的管理

空间数据元数据管理的理论和方法涉及到数据库和元数据两方面。由于

元数据的内容、形式的差异，元数据的管理与数据涉及的领域有关，它是通过建立在不同数据领域基础上的元数据管理系统实现的。国外在元数据方面已有许多成功经验，下面列举其中一例。

Francis P Bretherton（1994）等人认为元数据管理可通过元数据库（metadatabase）实现。如图 3－4 所示，在数据库管理系统的逻辑层上构筑元数据管理系统；元数据同数据一样，也存放在物理层中，但由一些软件通过一定的逻辑关系与逻辑层关联起来。与元数据管理系统直接相关的概念层用描述语言及模型定义了许多概念，比如，实体名称、别名、允许属性值的类型、缺省值、允许输出及输入的内容、临时实体的作用、元数据的变化和作模型等。通过这些概念及其限制特征，经过与逻辑层关联可获取、更新物理层的元数据及数据。

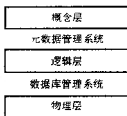


图3 4 元数据管理模式之一

3.3.2 地理空间数据元数据的概念和标准

1. 地理空间数据元数据及意义

上面介绍了一般的元数据概念，从本小节起，我们讨论地理空间数据的元数据。

地理空间数据元数据是说明地理空间数据的数据。在现代空间数据潮水般涌现和空间数据共享要求日益迫切的形式下，地理空间数据的海量、超海量特点和复杂性特点，使元数据对 GIS 领域的重要性，比别的领域更加突出。空间数据生产者和用户日益意识到管理和访问大型数据空间数据库和数据集的复杂性及很多面临的问题。数据生产者需要有效的数据管理和维护办法；用户需要找到更快、更全面和更有效的方法，以便发现、访问、获取和使用现势性强、精度高、易管理和易访问的地理空间数据。因此，空间数据的内容、质量、状况等元数据信息变得愈加重要，成为信息资源有效管理和应用的重要手段。

由于认识到元数据的重要价值，许多国家的中央政府、省和地方的部门、企事业和非盈利组织等已经对元数据引起高度兴趣。不少国家和国际性组织已经发布实施元数据内容标准。并开发了许多元数据操作工具。地理空间数据元数据标准和操作工具已经成为国家空间数据基础设施的一个重要组成部分，它也是“数字地球”的重要技术支撑条件之一。

元数据还有利于基于面向对象数据模型的地理空间数据管理（参照§3 1 4 和§3 2 4），可能成为这一极具前景的空间数据管理模式的重要组成部分。这是因为元数据有助于实现下述特点

- 完整性（completeness）。面向对象的地理信息系统和空间数据库的目标之一，是把事

物的有关数据都表示为类的形式，而这些类也包括类自身，即复杂的"类的类"结构。这就要求有支持类与类之间相互印证和操作的机制，而元数据可以帮助这个机制的实现。

• 可扩展性(extensibility)。有意地延伸一种计算机语言或者数据库特征的语义是很有用途的，如把跟踪或引擎信息的生成结果添加到操作请求中。通过动态改变元数据信息可以实现这种功能。

• 特殊化(specialization)。继承机制是靠动态连接操作请求和操作体来实现的，语言及数据库以结构化和语义信息的关联文件方式把操作请求传递给操作体，而这些信息可以通过元数据表达。

• 安全性(safety)。分类完好的语言和数据库都支持动态类型检测，类的信息表示为元数据，这样在系统运行时，可以被类检测者访问。

• 查错功能(debugging)。在查错时使用元数据信息，有助于检测可运行应用系统的解释和修改状态。

• 浏览功能(browsing)。为数据的控制类开发浏览器时，为显示数据，要求能解译数据的结构，而这些信息是以元数据来表达的。

• 程序生成(program generation)。如果允许访问元数据，则可以利用关于结构的信息自动生成程序，如数据库查询的优化处理和远程过程调用残体(stub)生成。

2. 地理空间数据元数据的常用具体概念

地理空间数据元数据的常用具体概念如下：

• 空间数据(geospatial data)。用于确定具有自然特征或者人工建筑特征的地理实体的地理位置、属性及其边界的信息；

• 类型(type)。在元数据标准中，数据类型指该数据能接收的值的类型；

• 对象(object)。对地理实体的部分或整体的数字表达；

• 实体类型(entity type)。对于具有相似地理特征的地理实体集合的定义和描述；

• 点(point)。用于位置确定的0维地理对象；

• 结点(node)。拓扑连接两个或多个链或环的一维对象；

• 标识点(label point)。显示地图或图表时用于特征标识的参考点；

• 线划(line)。一维对象的一般术语；

• 线段(line segment)。两个点之间的直线段；

• 线(string)。由相互连接的一系列线段组成的没有分支线段的序列，线可以自身或与其他线相切；

• 弧(arc)。由数学表达式确定的点集组成的弧状曲线；

• 链(link)。两个结点之间的拓扑关联；

• 链环(chain)。非相切线段或由结点区分的弧段构成的有方向无分支序列；

• 环(ring)。封闭状不相切链环或弧段序列；

- 多边形(polygon)　在二维平面中由封闭弧段包围的区域；
- 外多边形(universe polygon)。数据覆盖区域内最外侧的多边形，其面积是其他所有多边形的面积之和；
- 内部区域(interior area)。不包括其边界的区域；
- 格网(grid)。组成一规则或近似规则的棋盘状镶嵌表面的格网集合，或者组成一规则或近似规则的棋盘状镶嵌表面的点集合，
- 格网单元(grid cell)。表示格网最小可分要素的二维对象
- 矢量(vector)。有方向线的组合；
- 栅格(raster)。同一格网或数字影像的一个或多个叠加层；
- 像元(pixel)。二维图形要素，它是数字影像最小要素；
- 栅格对象(raster object)。一个或多个影像或格网，每一个影像或格网表示一个数据层，各层之间相应的格网单元或像元一致且相互套准；
- 图形(graph)。与预定义的限制规则一致的0维(node 点)、一维(link 或 chain)和二维(多边形)有拓扑相关的对象集；
- 数据层(layer)　集成到一起的面域分布空间数据集，它用于表示一个主体中的实体，或者有一公共属性或属性值的空间对象的联合；
- 层(stratum)。在有序系统中数据层、级别或梯度序列；
- 纬度(latitude)。地点位置向量与赤道平面之间的线面角；
- 经度(longitude)。经线面到格林尼治中央经线面的两面角；
- 经圈(meridian)。穿过地球两极的地球的大圆圈；
- 坐标(ordinate)。直角坐标或其他坐标；
- 投影(projection)。将地球球面坐标中的空间特征(集)转化到平面坐标体系时使用的数学转化方法；
- 投影参数(projection parameters)。对数据集进行投影操作时用于控制投影误差、变形实际分布的参考特征；
- 地图(map)。空间现象的空间表征，通常以平面图形表示；
- 现象(phenomenon)。事实、发生的事件、状态等；
- 分辨率(resolution)。由涉及到或使用的测量工具或分析方法能区分开的两个独立测量或计算的值的最小差异；
- 质量(quality)。数据符合一定使用要求的基本或独特的性质；
- 详述(explicit)。由对数或三个数分别直接描述水平位置和三维位置的方法；
- 介质(media)。用于记录、存储或传递数据的物理设备；
- 其他。

3 空间数据元数据的标准

伴随人类对数字地理信息重要性认识的加深，元数据标准化这一问题使逐渐成为共享地学信息的瓶颈之一。同一般数据相比，地理空间数据是一种结构比较复杂的数据类型，它既涉及到对于空间特征的描述，也涉及到对于属性特征以及它们之间关系的描述。因此，地理空间数据的元数据标准的建立比一般数据复杂；并且由于种种原因，某些数据组织或数据用户开发出来的，空间数据元数据标准很难为地学界所广泛接受。但是，空间数据元数据标准必须建立，这是空间数据标准化的前提和保证，只有建立起规范的空间数据元数据才能有效利用和共享空间数据。目前，空间数据元数据已形成了一些区域性或部门性的标准。表3-3列出了有关空间数据元数据的几个现有主要标准。

表3-3 现有的空间数据元数据标准（据G. G. Wilkinson）

元数据标准名称	建立标准的组织
CSDGM 地球空间数据元数据内容标准	FGDC，美国联邦空间数据委员会
GDDD 数据集描述方法	MEGRIN，欧洲地图事务组织
CGSB 空间数据集描述	CSC，加拿大标准委员会
CEN 地学信息－数据描述－元数据	CEN/TC287
DIF 目录交换格式	NASA
ISO 地理信息	ISO/TC 211

3.3.3 几种地理空间数据元数据标准及实例

为了帮助理解空间数据元数据概念，这里选取几种有一定代表性的空间数据元数据标准及实例加以说明。

1. 美国 FGDC 元数据标准

美国联邦空间数据委员会（Federal Geographical Data Committee，简称 FGDC）的地理空间数据元数据标准是影响较大的标准之一。该标准用于确定地理空间数据集的元数据内容，于1992年7月开始起草，于1994年通过，并发布该标准的第一版，1997年完成第二版。

表3-4 CSDGM 子集一览表

主要子集	次要子集
标识信息	引用文献信息
数据质量信息	时间信息
地理空间数据组织信息	联系信息
地理空间参考系统信息	
实体及属性信息	
发行信息	
元数据参考信息	

FGDC 的地理空间数据的元数据内容标准(Content Standards for Digital Geographic Metadata,简称 CSDGM)说明一组地理空间数据的元数据的信息内容,提供与元数据有关的术语和定义,说明哪些元数据是必需的、可选的、重复出现的,或者是按 CSDGM 产生规则编码的。CSDGM 是参照文件,它说明当用户在评价数据集的用途、获得该数据或有效使用数据时,需要知道的信息。

第二版 CSDGM 打印文本有 83 页。包含 7 个主要子集和 3 个次要子集(表 3-4),共有 460 个像上一小节所例举的那种元数据实例(含复合实体)和元数据元素。元数据元素是元数据的关键术语,是其最基本的单元。一个元数据元素说明地理空间数据的某一方面特征。按数据库语言,它们是填入数据的"字段"。一个或若干个元数据元素组成元数据实体。复合实体则由元数据实体、元数据元素和/或其他复合实体构成。每个元数据元素、实体或复合实体均需说明其名称、定义、类型、值域和简称等特征信息。元数据子集是由若干元素简单的或复合的元数据实体组成的集合。

CSDGM 标准规定了三种性质的子集、实体和元素。第一种是必需的,即必需提供的信息。第二种是一定条件下必需的,即如果正在建立的元数据包含某子集、某个实体,或某个元素说明的特征,则必需提供的信息。第三种是可选的,即该信息是可选的,由用户决定是否将其包含在元数据文件中。

CSDGM 还可以用图示方法表示元素。当用图示方法表示时,每个元数据元素为一个小图框,框内填写元素的关键字名。实体及复合实体则表示为围绕数据元素和/或其他元素的方框。图 3-5 是一个复合实体"数据集范围"的框图,它由地理坐标、时间范围和高程范围等三个实体组成。其中,地理坐标为复合实体,它由四个元素和一个"地理区域"实体组成。地理区域实体又由两个元素组成;时间范围实体由四个元素组成;高程范围实体则由三个元素组成。

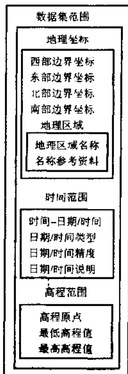


图 3-5　复合实体数据集

元数据子集、实体和元素之间的关系也可以用通用建模语言(UML)描述。

在元数据文本文件中,这种子集、实体和元素关系,用元素比实体缩进一格的办法表示(称为分层缩排),或者用编号,或者在编号后增加一个小数点和下一层编号表示。例如,实体 3.1 可以有元素 3.1.1 和 3.1.2 等。以下是美国地质测量局(USGS)发布的 30′×30′数字高程模型(DEM)元数据中数据志(lineage)子集的部分内容,它明显地体现了缩排结构(其中,黑体部分是元数据数据值,其余是元数据子集、实体和元素名称)。

Lineage:

　　Source_Information:

　　　　Source Citation:

Citation Information:

Originator: **U. S. Geological Survey**

Publication Date:

Title: **Digital contour lines**

Geospatial Data Presentation Form **map**

Publication Information:

  Publication Place: **Reston, VA**

  Publisher: **U. S. Geological Survey**

  ....

这里有几点需要注意:

第一,元数据子集、实体和元素名称必需与标准完全一致。

第二,元数据的实体(含复合实体)与其元素之间的关系必须处理好,即元素总是跟随各自的实体。

第三,元数据即使符合标准,看起来也可能是不一样的。这是由于标准仅说明元数据的内容,而不是它的格式。不同格式包括:元数据元素分层缩排、编号系统、元素名后加冒号、在元素名称下划线,或者将其作为分开的单词等。可以利用元数据操作工具软件将它们规则排列起来。操作工具可以读入元数据文件,输出与标准一致的元素名称。

美国 FGDC 元数据标准除在美国国内广泛使用外,加拿大、印度等国也已采用,并作为各自的国家标准。下面即将介绍的 ISO/TC 211 正在利用该标准文本作为基础,制订相应的国际标准。

2. ISO/TC 211 的元数据标准草案

ISO/TC 211(国际标准化组织/地理信息业技术委员会)自 1994 年成立以来,就将地理信息元数据作为其首批研制的 20 个国际标准之一。几年来,参加该项标准研制的第三工作组专家,经过艰苦的努力,于 1996 年 3 月完成第一版工作草案,先后完成近 10 个更新版本。该标准以 FGDC 等现有标准为基础,按照国际标准化组织制订的标准规则要求制订。其工作范围是:定义说明地理信息和服务所需要的信息。它提供有关地理数据标识、覆盖范围、质量、空间和时间模式、空间参照系统、发行等信息。该标准适用于数据集编目、数据交换网络,以及数据集的详尽说明。适用于地理数据集、数据集系列、地理要素和属性。它定义:

- 必选和一定条件下必选的元数据实体和元素,即提供元数据应用、数据发现、决定数据对于应用的适合程度、数据访问、数据转换和数据应用所需要的最少的核心内容;
- 可选的元数据元素——如果需要,对地理数据的更加详细的标准说明;
- 为适应特殊需要对元数据进行扩展的方法。

尽管该标准是用于数字数据的,它的原则也能扩展用于许多其他形式的地理资料,如地图、图表、文本文件等。

该标准确定了两级元数据,即:

一级元数据——编目信息,包含数据集编目所需的最少的元数据内容。

二级元数据,它包含八个子集和三个可重复的实体。八个子集是:

- 标识信息。包含唯一标识一个数据集、说明其空间和时间范围、状况、法律限制和保

密限定所需的信息；

- 质量信息 包含数据集质量的一般评价；
- 数据志信息 包含有关数据集应用、数据源，以及生产数据集时所用的工艺方法；
- 空间数据表示信息 包含与数据集中表示空间信息所用方法有关的信息；
- 参照系统信息 包含数据集中应用的空间和时间参照系统说明；
- 要素分类信息。包含数据集中具有的要素类型、要素功能、要素属性和要素关系的定义和说明；
- 发行信息 包含有关获取该信息所需的数据发行者及买卖权限的信息；
- 元数据参考信息 包含元数据现势性及其负责单位信息

三个可重复的实体不是独立的元数据子集，它们不单独使用，而总是作为一个单元插在前述某一个子集中，且仅作为该子集的实体和元素 这三个可重复的实体是：

- 引用文献信息实体。提供引用文献的标准格式；
- 负责单位信息实体 包含与数据集有关的单位和/或个人的标识；
- 地址信息实体。提供与数据集有关的单位和/或个人的地址及其他通信办法。

标准定义了每个元数据子集、实体和元素的八个特征，即名称、标识码、定义、性质、条件、最大出现次数、数据类型和值域。表3-5是一级元数据 编目信息的几个元素的特征表实例（其中，标识码略）。

表3-5 一级元数据元素特征实例表

序号	名称	定 义	性质	最大出现次数	数据类型	值域
1	编目信息	唯一标识完整数据集和、或数据集系列的信息	M	1	元数据子集	序号2—55
2	名称	数据集名称	M	1	文本	任意长文字
3	版本	数据集版本	0	1	文本	任意长文字
4	系列名	数据集所属的系列产品总称	0	1	文本	任意长文字
5	出版标识	数据集所属的系列新产品出版标识信息	0	1	文本	任意长文字

目前，该标准工作草案一级元数据包含55个元数据实体和元素，二级元数据的八个子集和三个可重复使用的实体共包括452个实体和元素。

考虑到不同数据集、数据集系列、要素和属性的不同特点，工作草案不但定义了子集、实体和元素的三种性质，即"必选"、"在一定条件下必选"和"可选"三种性质，而且提供了扩展的方法和摘录其部分内容的原则与方法。

该标准是迄今最为完整，也最为复杂的标准方案。它的制订，必将影响许多国家的地理信息元数据标准的制订和实施，并进而对全球数据管理与服务产生很大影响。

3. 中国基础地理信息元数据实例

同国际上一样，中国也有自己的地理空间数据元数据标准。下面列出国家基础地理信息系统全国1∶400万地形数据库元数据（一级，说明性元数据）的所有元素名录 为节省篇幅，略去一部分填写的内容；且一部分本应各占一行的细目合起来放在一个括号中。

编目信息

数据集中文全称：国家基础地理信息系统全国1∶400万数据库
数据集中文简称：1∶400万数据库
数据集英文全称：1∶4M Data Base of the National Fundamental Geographic Info
数据集英文简称：1∶4M DB(NFGIS,PRC)
版本：1.0
系列名：国家基础地理信息系统
出版系列标识
出版日期
数据集所属项目标识信息：
项目名称：国家基础地理信息系统网络体系
项目类型：部门攻关项目
负责单位信息：
（负责单位名、负责人姓名、负责单位作用、国别、政区、城市、地址、邮政编码、网址、电子邮件地址、电话号码、传真号码）
数据范围：
地理坐标：
（西部、东部、北部、南部边界坐标、地理区域名称）
时间范围
比例尺：1∶400万
分辨率：
数据集语言：中、英文
数据集内容信息：
（摘要、目的、进展、专题名称、关键词、限制信息、访问限制、使用限制）
数据志说明：
质量说明：符合1∶100万地形图精度
数据表示类型：矢量
数据项：
空间参照系统类型：经纬度
发行信息：
（发行单位名称、发行格式、发行介质、数据量、网上发行地址、浏览图网址、定价）
元数据参考信息：
（元数据级别、元数据负责单位、元数据负责人、政区、城市、地址、邮政编码、电话号码）

3.3.4 空间数据元数据的获取和应用

空间数据的空间位置特征和属性特征要求对数据的各种操作，从数据获取、数据处理、数据存贮、数据分析、数据更新等方面应有一套面向地理对象的方法，相应的空间数据元数据的内容及相关的操作也就具有了不同于其他种类数据元数据的特点。

1 空间数据元数据的获取过程

空间数据元数据的获取是个较复杂的过程，相对于基础数据的形成时间，它的获取可分为三个阶段：数据收集前、数据收集中和数据收集后。对于模型元数据，这三个阶段分别是模型形成前、模型形成中和模型形成后。

第一阶段的元数据是根据要建设的数据库的内容而设计的元数据，内容包括：普通元数据，例如，数据类型、数据覆盖范围、使用仪器描述、数据变量表达、数据收集方法等；以及专指性元数据，即针对要收集的特定数据（如中国1950—1980年30年间的逐旬降水数据）的元数据，如数据采样方法、数据覆盖的区域范围、数据表达的内容、数据时间、数据时间间隔、空间上数据的高度（或深度）、使用的仪器、数据潜在利用等。

第二阶段的元数据随数据的形成同步产生，例如，在测量海洋要素数据时，测点的水平和垂直位置、深度、温度、盐度、流速、海流流向、表面风速、仪器设置等是同时得到的。

第三阶段的元数据是在上述数据收集到以后根据需要产生的。它们包括：数据处理过程描述、数据的利用情况、数据质量评估、浏览文件的形成、拓扑关系、影像数据的指示体及指标、数据集大小、数据存放路径等。

2. 空间数据元数据的获取方法

空间数据元数据的获取方法主要有五种：键盘输入法、关联表法、测量法、计算法和推理法。键盘输入一般工作量大且易出错，如有可能应尽量避免，但对某些元数据而言（如数据变量表达的内容）只能由键盘输入。关联表法是通过公共项（字段）从已存在的元数据或数据中获取有关的元数据。例如，通过区域的名称从数据库中得到区域的空间位置坐标等。测量法容易使用且出错较少，如用全球定位系统（GPS）测量数据空间点的位置等。计算法指由其他元数据或数据计算得到的元数据。例如，水平位置可由仪器设置及时间计算得到，区域的面积可由多边形拓扑关系计算出来。该方法一般用于获取数量较大的元数据。推理法指根据数据的特征获取元数据。

在元数据获取的不同阶段，使用的方法也有差异。在第一阶段主要是键入法和关联表法；第二阶段主要采用测量法；第三阶段主要方法是计算法和参考方法。

3. 空间数据元数据的应用及其意义

前面介绍元数据概念时曾提到过元数据的应用意义，这里进一步说明。

首先，元数据可以帮助用户获取数据。信息社会的一个特点是数据（信息）通过互联网等各种方式实现全球、区域或部门的共享，而缺乏符合区域、部门和行业标准的元数据，往往是实现数据共享的瓶颈。一个完整的地学数据库除应提供空间数据和属性数据外，还应提供丰富的引导信息，以及由实际数据得到

的分析、综述、索引等信息。这些信息能帮助用户明白诸如"这些数据是什么数据"、"这个数据库对我是否有用"和"怎样得到这些数据"等一系列问题,从而能有效地获取所需数据,实现数据共享的目的。

其次,可利用元数据进行空间数据质量控制。地理空间数据一般都存在数据质量问题。影响空间数据精度的原因主要有两个方面:一是源数据本身的精度;二是数据加工、处理过程对精度质量的影响。无论在哪一方面,元数据可用来帮助空间数据的质量控制。例如,给出有准确定义的数据字典,以说明数据的组成,各部分的名称和表征的内容等;保证数据逻辑科学地集成,如植被数据库中不同亚类的区域组合成大类区,这要求数据按一定逻辑关系有效的组合;给出说明数据来源、数据的加工处理工程、数据释译的足够信息等。这类元数据往往需要地学和计算机双重领域背景。例如数据逻辑关系和编码系统等需要地学背景;而数据质量的控制、提高和数据再生产需要数据输入、查错、处理等专业背景。

第三,元数据在数据集成中的应用。数据集层次的元数据记录数据格式、空间坐标体系、数据的表达形式、数据类型等信息;系统层次和应用层次的元数据则记录数据使用软硬件环境、数据使用规范、数据标准等信息。这些信息在数据集成的一系列处理中,如数据空间匹配、属性一致化处理、数据在各平台之间的转换使用等是必需的。这些信息能够使系统有效地控制系统中的数据流。

第四,元数据在数据管理中的应用。元数据在数据库管理中的应用,主要表现在可以借助元数据高效地建立并维护符合一定标准和规范的科学数据库。通过元数据建立的逻辑数据索引,可以高效查询检索分布式数据库中任何物理存储的数据,避免数据的重复存储,缩短数据库建库及更新时间,实现系统资源的合理利用和分配。同时,元数据的开发和利用,也大大增强数据库的功能(如数据转换、数据分析等)。而且,应用元数据,数据库管理人员就很清楚各数据的来源、数据的精度、数据的期限或寿命、数据的物理存储地址、各数据间的关系等情况,从而知道哪些是需重点维护的数据,以及如何维护这些数据,达到减少数据库维护的工作量、降低数据库的维护费用目的。

第四章
地理空间数据处理

§4.1 有关地理空间数据处理的总体说明

4.1.1 地理空间数据处理总体说明的必要性及内容

1. 地理空间数据处理总体说明的必要性

数据处理(Data processing)是现代信息科学或计算机科学中应用得最广泛、随意性最强的用语之一,它没有统一的、确定的定义。数据处理的含义可以十分广阔,信息系统或某些计算机系统的大多数功能都可以视为"数据处理"。例如,国际上某些重要词典在简释"信息系统"一语时,就说"信息系统是一种数据处理的系统",即信息系统的功能可以用"数据处理"来概括。另一方面,"数据处理"一语也常在狭义上使用,专指某些特定的数据处理功能,这种理解不仅因行业,而且因人而异。

地理信息科学领域的情形也一样。从广义上说,地理空间数据从采集到输出,整个流程中的绝大部分工作都可以看作是进行数据处理,"地理空间数据处理"可以涵盖GIS中大部分内容。但实际上,正是因为"数据处理"涵盖太广、太泛,人们必须具体地区分和研究"数据处理"功能的不同部分,将"数据处理"功能分门别类,如数据采集、检验、编辑、格式化、转换、概括、存储、组织、分析和显示等。在这种分类中,有些学者又把其中一部分功能称为"数据处理"功能。这样,就有了"数据处理"的狭义理解。例如,我们在§1.1.4.2中,就把检验与编辑、格式化、转换和概括称为数据处理。

但是,在"地理空间数据处理"的狭义理解和使用上,不同人的差别是比较大的。正因此,在国内外各种GIS教材或书籍中,"数据处理"以外的其他章节,如数据结构和空间分析等章节,内容甚至名称,差异都不大,但涉及"数据处理"的章节的内容和名称,就特别不一致:少数书籍有地理空间数据处理的专门章节,但涵盖各不相同;部分书籍将数据处理与某些作者认为难以分割的内容合

为一个章节，如“空间数据采集与处理”或“空间数据处理与分析”等；还有的GIS书籍将所有的“数据处理”内容分散到不同的章节，以致全书没有一个以空间数据处理命名的章节。

各种章节安排都有各自的合理性。

尽管如此，有一点是共同的：大家在行文中都大量使用“数据处理”一语，但没有什么书籍对“地理空间数据处理”这一用语的涵盖进行讨论和交代。

本书含有“地理空间数据处理”一章，并在此专门对地理空间数据处理及其所涉及概念作一总体说明，主要有下述原因：

第一，同其他GIS书籍一样，本书中，也有不少可以看作是“地理空间数据处理”的内容被分散到其他章节；但是，确实还有一些“空间数据处理”的内容不适宜纳入别的章节，需要专辟一章；而这一章又找不到比“地理空间数据处理”更好的名称。可见，本章有一点“不得已”的性质。既然如此，我们就有必要交代这里的“来龙去脉”，告诉学生：本书认为可以看作是“地理空间数据处理”的内容有哪些？其中，哪些为本章内容？哪些在本书中被安排到前面或后面的哪些章节？哪些内容只需略加交代，或主要是其他课程的内容？并简述原因，使读者对本章学习心中有数。

第二，这种说明不仅是本章学习的需要，而且对理解本课程整个知识体系也有助益。这是因为，广义上的“地理空间数据处理”可以涵盖地理信息系统的大部分内容，交代有关“空间数据处理”的来龙去脉，涉及到前面或后面章节内容的说明。而本书发展到现在，适逢课程中期（期中考试前后），在这时，对“空间数据处理”有关概念作一次广泛的整理，即进行一次跨章节内容的框架性说明，对全书的学习有较好的承上启下的作用：一方面能复习前面的内容，另一方面有助于今后的学习。

第三，在进行框架性说明之时，我们还可以趁便把GIS特色和非GIS特色的“地理空间数据处理”区分开来。事实上，由于地理空间数据的特殊复杂性，地理空间数据处理功能，既有GIS特色功能，也势必包括一般计算机图形图像处理和数据库处理等功能。一般GIS书籍不区分这两类功能。我们从多年的GIS教学实践经验感觉到，如果能加以区分，强调哪些是GIS特有的数据处理功能，哪些是GIS所使用的一般计算机功能，也会对读者理解地理信息系统基础知识框架有助益。

2. 本节的内容和做法

我们将采取下述做法对地理空间数据处理的来龙去脉进行总体说明。

第一，尽可能完全地“网罗”可能看作是“地理空间数据处理”的内容或概念，形成一个广义上的“地理空间数据处理”内容或概念的集合。

第二，对上述集合中的各种“地理空间数据处理”内容或概念进行梳理，并

分门别类,区分出非 GIS 特色的地理空间数据处理(包括常规数据处理、一般的图形和图像处理等)和 GIS 特色的数据处理功能;说明本书所认定的狭义 GIS 数据处理功能;最终形成一个较完整的地理空间数据处理之概念框架,并清晰地归纳成一张表,即表 4 - 1。

第二、分别就表 4 - 1 的每一条数据处理内容进行具体说明。如果该内容本书业已讲解过,或者将放在后面的章节中讲述,我们将说明在哪一个章节,以及为什么要那样做的原因;有些内容将在本章用专门篇幅讲授,我们也将加以说明;还有些本书不拟用专门篇幅讲授的内容,我们也将在本节中顺便加以简单解释。此外,我们还要说明各种地理空间处理内容之间的关联。由此亦可见,这样做实际上就是将本课程大部分内容进行一番梳理,是一种承上启下的工作。

3. 广义的地理空间数据处理的概念框架

按上述要求建立的广义地理空间数据处理的概念框架,如表 4 - 1 所示。图中将数据处理类别分为 5 大类。其中,第 5 类为一些从广义上才能看作是地理空间数据处理的功能;1 ~ 4 类为本书所认定的狭义 GIS 数据处理功能。后者中,只有第 4 类为 GIS 特色的空间数据处理功能;1 ~ 3 类为非 GIS 特色的空间数据处理,在地理信息系统中,这些功能也经常被用来处理地理空间数据,但这些功能本身并非 GIS 特色。

下面就来按表 4 - 1 的顺序对各种地理空间处理一一进行评述。

表 4 - 1 广义上的地理空间数据处理的概念框架

空间数据处理类别	数据处理功能	简单描述
1 普通数据处理	单纯属性数据变换及处理	GIS 在普通属性数据库中的查询、选择、增补、提取、关联和逻辑的、数学的计算等
2. 一般的图形处理(CAD 等图形处理软件也具有的数据处理功能)	几何计算及查询	距离、长度、面积等的计算及查询
	一般图形编辑	点、线和多边形的增删、修改、移位、恢复和复制等,以及点、线和多边形之间某些相互关系的修改,如两线端点咬合、端点与线咬合等
	线状对象其他操作	线相交、曲线光滑处理和平行线处理等
	多边形其他基本操作	点在多边形内、线与多边形相交和多边形相交判别,区域填充等
	图形窗口裁剪与合并	使用规则窗口或不规则多边形模板的图形裁剪,和多边形等图形或图形区域的合并
	图幅的几何接边	相邻图幅拼接时边界两侧线和多边形的几何衔接
	图层(几何)叠加	多重图层叠加显示,即不同图层进行几何的、机械的叠加
	坐标变换	地图投影以外的坐标变换
	格式转换	不同软件开发商或数据供应商的不同数据格式之间的转换
	三维处理	以规则形体为主要对象的三维立体表达和变换

续表

空间数据处理类别	数据处理功能	简单描述
3 其他非GIS特色的数据处理	一般的图像处理	图像显示、编辑、增强和变换等
	其他多媒体技术处理	GIS所用到的视频、声频等处理
4 GIS特色的数据处理	属性与图形交互编辑	基于空间对象唯一标识符，图形编辑中进行相应属性数据的编辑，或属性数据的编辑涉及到相应图形编辑
	GIS图幅接边	不仅进行几何接边，还要进行逻辑接边，即同一地物在边界两侧的属性必须一致
	拓扑关系生成和编辑净化(cleaning)	找交点形成结点和线，装配多边形、清除假结点、点在多边形内的判别等
	制图概括(或制图综合)及有关编辑	根据地图比例尺、用途、地理特征等条件，进行地图内容取舍、符号化、地物移位和制图对象概括和归纳推理等，以及专为制图概括所进行的编辑，如曲线化简等
	空间数据质量与精度方面的处理	进行空间数据质量的控制、总结和评价等，以保证空间数据较好的准确性、一致性和完整性，以及它们三者之间统一性
	地理坐标转换和地图投影	其他坐标向地理坐标的转换、不同地理坐标之间的转换、地表曲面上的地理坐标向平面上投影、精校正、坐标配准等
	空间数据内插	通过已知点或分区数据推求其他地点的可能取值
	三维立体处理	面向不规则地表的GIS特色的三维模型和显示等处理
	地图输出前的预处理	地图符号、注记、图例等GIS特色制图处理
	矢量、栅格数据转换	涉及地理属性数据挂联的矢量、栅格数据的相互转换
5 广义上的空间数据处理	数据采集和输入	从现实世界的采集、观测，以及从现存文件、地图中获取地理空间数据，并输入到计算机中
	显示与输出	通过屏幕等多种途径，可视化地表达地理空间数据和信息
	数据存储、组织与管理	按GIS表达和运作的要求来组织和管理地理空间数据，涉及空间数据结构和GIS数据库系统技术
	空间查询和分析	位置和属性交互的空间查询、已形成某些GIS软件中的固有功能的程式化空间分析和其他空间模型分析等

4.1.2 非 GIS 特色的地理空间数据处理

1. 普通数据处理

在地理信息系统运作中，有时需要单纯在属性数据的范围内进行变换或处理，而暂时不涉及空间位置几何数据的处理，这时只需要采用一般的（关系）数据库技术及其处理功能就够了。举一个普通的例子：在 §1.4.1 中我们曾列出地理信息系统的五项主要任务，其中第一条是"某个地方有什么属性"的问题。解决这个问题，如果不是从地图图元或图形出发（例如通过鼠标点击地物的图形），而是从地名或邮政编码出发来查询属性，那就只需要在属性数据的范围内进行处理就行了。例如，查询地名为"宁波"的城市的某些属性，那就完全可以仅在常规数据库范围内，通过其查询功能来解决问题。可见，这种单纯的属性数据处理或变换也是地理空间数据处理的一部分。

GIS 中只需要单纯处理属性数据的情形还有很多。例如，通过 GIS 的属性数据库来进行组合条件的地理查询，或进行多因子分析评价计算（如评价结果 $=\Sigma$ 权重因子 × 某属性）等。对于单纯的属性数据处理，常规数据库的数据处理功能，如查询、选择、增补、提取和关联，以及逻辑和数学计算等，已足够强大了。这时，完全没有必要进入 GIS 软件的图形处理模块。例如，采用 Arc/Info 软件平台的用户都有体会：单纯进行属性数据处理的最便捷途径，就是利用 Arc/Info 的常规数据库模块（tables）来操作，而不必进入图形/属性交互编辑模块（ArcEdit）。

2. GIS 中的一般的图形处理

表 4－1 中的第二类"一般图形处理"，指 CAD 等图形处理软件也具有的数据处理功能。它们虽不是 GIS 所特有的，但是，在 GIS 数据处理中也相当常见。下面按表 4－1 的顺序对这一类进行逐条说明。

"几何计算及查询"是一般图形处理中最普通的一种，指点、线和多边形各种有关几何特征的计算。GIS 和图形软件都需要进行几何计算。最基本的几何计算有三种：点与点之间的距离、线的长度和多边形面积的计算。几乎所有的软件都有这几种默认的功能，计算的结果都能及时显示和查询。

"一般图形编辑"指点、线和多边形的增删、修改、移位、恢复和复制等，以及点、线和多边形之间某些相互关系的修改，如两线端点咬合、端点与线咬合等，这些都是 GIS 所不可缺少的。"线状对象其他操作"和"多边形其他基本操作"（表 4－1 中接下两栏）实际上也是图形编辑功能。表中所给出的各种具体操作例子，从名称容易推想其含义。例如"线相交"，很容易想到它就是 GIS 和图形软件中常见的、确定线与线交点的功能。"区域填充"，又称为多边形填充，在 §2.4.2.2已讲过。除了那里讲过的内容外，它在矢量数据输出中也很有用，借

助于它可以用不同图案或色彩来显示或输出不同的区域多边形。

“图形窗口裁剪与合并”，指按用户指定的窗口或图框，从已有图中裁剪出图的局部；窗口或图框可以是规则形状的，如矩形等，也可以是用户点取的任意多边形。关于这种数据处理，我们还将在§5.1.2.4中结合“几何窗口查询”进一步解释。

“图幅的几何接边”，指相邻图幅拼接时相邻边两侧的线和多边形的几何衔接。由于原图或数字化过程中的多种原因，相邻图幅的电子地图在它们的边界处不可能完全吻合，跨图幅的地物线在边界处一般都有点“错开”。几何接边的主要任务，就是使错开的地物线吻合。

“图层(几何)叠加”，指不同图层进行几何的、机械的叠加，即进行多重图层叠加显示。这种处理不涉及数据库的属性挂联，一般也不改变各图层数据文件的本身。例如，将红色的道路图层与蓝色的水系图层叠加起来显示，呈现既有红色道路，又有蓝色水系的图层。

“坐标变换”，这里指地图投影以外的普通坐标变换。它也是一般图形软件的常见功能。GIS运作也常常进行普通坐标变换。例如，经扫描矢量化或手扶跟踪数字化仪输入的地图，通常采用以英寸为单位的普通直角坐标系；在未注册到大地坐标系以前进行的坐标变换，都是普通的坐标变换。

“格式转换”，即不同软件开发商或数据供应商的不同数据格式之间的转换，包括常规意义上的矢量、栅格数据的转换，也是一般图形处理都会遇到的问题，并非GIS特色。

最后，CAD等图形软件也一般都具有三维立体处理功能，有的三维功能还非常强大。例如流行的AutoCAD软件不仅本身有三维功能，它所属的公司Autodesk的另一产品3D Studio VIZ的三维模拟、动画飞行等功能，达到令人赞叹的地步。因此，三维处理并非GIS特色。当然，CAD类的三维功能与其本身一样，面向工程设计，因而擅长于规则形体(一般可用数学解析表达)的三维模拟和表达。

应当强调，GIS和一般图形软件虽然有不少相同或相似的图形处理功能，但是，这些处理功能在GIS和图形软件中的作用常常是不同的。例如，“线相交”和“端点咬合”等在GIS中，是建立拓扑矢量数据结构的必要条件(参见§2.2.4.4)；而一般图形软件虽有此功能，但并不一定施行，因为很多图形软件面向工程设计，只要在绘图和显示时，看上去是相交和连接的就行了。

又例如，“图层叠加”在CAD等图形软件中的主要作用是复合显示，即看上去的叠加。如上述红色道路和蓝色水系叠加，看上去合而为一图层，实际上存储器中的道路层和水系层数据文件并无改变，只是被调出来进行屏幕显示而已。GIS虽然也常进行这种单纯的图层叠加显示，但它还有更进一步的、特色的

图层叠加处理，在那种叠加中，不同图层的点、线、面真正结合成一层，重新找交点，分配属性，重新建立拓扑关系，形成新的文件（那种处理称为空间叠置分析，详见§5.2.2）。

GIS的三维功能也与CAD类有所不同，GIS更擅长于进行不规则自然形体的三维模拟和表达，并且更看重在三维环境中的空间分析。

最后值得一提的是，由于地理空间数据的特殊复杂性，GIS领域在格式转换上的麻烦比一般图形处理软件更多一些，面临更严峻的专业软件集成（互操作性）和数据集成的问题

3. GIS中的一般的图像处理

表4-1中的第二大类“其他非GIS特色的数据处理”，主要包括一般的图像处理和多媒体技术处理。现代GIS系统大量地采用办公自动化和多媒体技术。因此，用多媒体技术处理GIS数据也成为地理空间数据处理的一种成分但多媒体数据处理不是本课程内容。本小节主要谈与GIS有关的一般图像处理问题。

我们已多次讲过，与图像处理技术最密切的，是栅格型GIS。就目前而言，栅格型GIS的主要研发和应用领域有三个：遥感图像处理、栅格型数据的多因子综合分析和三维立体处理。其中，遥感图像处理内容最丰富，涉及一般图像处理技术也最多。按照GIS专业一般的课程划分习惯，“地理信息系统”课程在栅格数据结构的基础上，通常只讲栅格型数据多因子综合分析和三维立体处理的内容（在本书中分别见§5.3.1和本章第4节）；而“遥感图像处理”通常单独作为一门课程讲授。

由于GIS课程一般不展开遥感图像处理的内容，而GIS特色的图像处理与一般图像处理的区别，又主要体现在遥感图像处理领域中，为此，我们在这里顺便介绍一下“遥感图像处理”的内容及其与一般图像处理技术的关系。“遥感图像处理”一般包括六个基本部分：

(1) 遥感和遥感图像基本知识；

(2) 图像及显示基础知识；

(3) 图像增强和有关变换；

(4) 遥感图像分类；

(5) 遥感图像校正（含地理坐标变换）；

(6) 栅格型GIS及空间分析方法（仅部分遥感处理课程或书籍有此部分）

在上述部分中，第二和第三两部分，大体就是“一般图像处理”的内容 其中，图像及显示基础知识主要讲解计算机图像文件及其格式、存储和显示原理，包括红绿蓝（RGB）调色、真假彩色显示和文件编辑等；图像增强和有关变换则主要讲授通过各种变换方法改善图像的色调、反差、纹理、局部特征和清晰度

等，以便使图像变得更好辨认、判读或鉴赏。这些都是一般图像处理很基本的内容。事实上，如果打开某种通用型图像处理软件(简单的如 Windows 下的"图像处理"，功能较强的如 Photoshop 等)，很容易在软件的各级菜单中发现这些图像处理功能。

GIS 特色的图像处理内容，除有关遥感的基本概念外，主要是遥感图像分类和遥感图像校正两部分。前者通过多种手段或途径从多波段遥感图像中提取各种各样的分类信息(参见 §2.3 和 §2.4.2.3 有关内容)，即通过图像处理判别区域内有什么样的土地、植被、生态、污染和灾害等，并尽可能提高分类或判别的精度或正确率。遥感图像校正则主要是从辐射上、几何上保证图像的精度和数据质量，保证图像准确地落实到大地坐标系中，进而能顺利地与其他来源的 GIS 数据源进行集成分析。

遥感图像处理与一般图像处理有鲜明的区别。其中主要一点是：遥感图像处理是一种科学的图像处理，它求"真"求"是"；而一般的图像处理则偏重大众的需求，以观赏为主要目的，为了求"美"，它可以不求真实，甚至刻意修饰。例如，在肖像照片的扫描图像上，修饰眉毛和嘴巴，改善面庞的颜色等。

从 20 世纪 70 年代出现多波段遥感图像起，遥感图像处理技术一开始就立足于计算机方法研究，而且得到很多其他有关领域难得的巨大人力、物力和财力的投入，经过近 30 年世界各国大量专家学者的努力，现已发展成为一个具有丰富内容的学科，为现代海量数据自动信息提取技术做出了重要的贡献。

4.1.3 GIS 特色的数据处理和广义的空间数据处理

本节包括表 4-1 的第四、第五两大类的内容。其中，第四大类"GIS 特色的数据处理"将分为三个部分(即前面三小节)来讲述。

1. GIS 特色的数据编辑

GIS 数据编辑的最大特色是属性与图形交互编辑。这种编辑基于空间对象唯一标识符，要求保持空间位置数据和属性数据二者编辑修改的同步性和一致性。Arc/Info 的图形/属性编辑模块(ArcEdit)就是一个典型的 GIS 数据编辑环境。在其中，用户可以方便地交互进行图形编辑及相应属性数据的编辑，保证修改过的点、线和面空间对象能及时赋予修正的属性。GIS 的图幅接边也是一样，不仅要进行几何接边，把两幅图上几何上错开的地物线条接上，还要进行逻辑接边，即保持同一地物在边界两侧属性的一致。这也要求有属性与图形交互编辑的环境。同图形/属性编辑模块一样，GIS 软件的图幅接边模块一般也都提供一系列自动和半自动相结合的交互编辑手段。

另一种典型的 GIS 特色的数据编辑是拓扑编辑。拓扑编辑处理的内容通常包括：其一，在地理空间数据中建立或生成拓扑关系，其步骤是：找交点定义

结点和线、装配多边形、建立多边形标签、建立线属性表和多边形属性表等。其二,进行数据净化(cleaning),如清除假结点和多余的中间点等。其三,拓扑编辑,按拓扑数据的要求进一步进行编辑。拓扑编辑是比较费工费时的,这是获得有较优性能的拓扑数据的代价。§2.2.2.4和§2.2.4.4已比较详细地讲解过这方面的道理。

还有一种GIS特色的数据编辑是计算机制图概括(或制图综合)。制图概括指根据地图比例尺、用途和地理特征等条件,进行地图内容的取舍、符号化、地物移位和制图对象概括和归纳推理等。如地图学所述,制图概括很大程度地依赖于地图学家的专业智能和人为处理。因此,在计算机中实施制图概括的难度较大,需要尽可能采用计算机人工智能技术。目前在GIS软件中,具有一定自动化程度的制图概括功能有:线条的简化、平滑;(比例尺缩小)缩编地图时部分实现地物数量、形态等的筛选和简化表达;以及在屏幕上缩放地图时,自动地减少或增加图层,以保持适当的图面负载量等。总的来说,目前GIS制图概括自动化处理的能力还不强,GIS中的大多数制图概括工作,仍然需要在GIS特色的编辑环境下,由制图专业人员通过人工干预方式在屏幕上手工操作完成。

2. 空间数据质量与精度控制

数据质量和精度问题一直是GIS最重要的基础研究领域之一。考虑到本书属本科生学习层次,我们仅在此做一些介绍。有兴趣进一步了解此领域的读者可参阅陈述彭、鲁学军、周成虎的《地理信息系统导论》。

数据质量是空间数据在表达位置、属性和时态特征三个基本要素时,所能够达到的准确性、一致性和完整性,以及它们三者之间统一性的程度。数据质量是数据整体性能的综合体现。表达数据质量的参数有:误差(error,与真值的差异)、准确度(accuracy,与真值的接近程度)和精密度(precision或resolution,数据表达的精密程度)等。准确度和精密度有时合称为精确度,简称精度。

不同科学领域有不同的质量标准。GIS领域的质量标准通常称为空间数据质量标准。空间数据质量标准是生产、使用和评价空间数据质量的依据。目前,世界上已经建立了一些数据质量标准,如美国FGDC的数据质量标准等。空间数据质量标准包括下列要素:数据情况说明、位置精度或称定位精度、属性精度、时间精度、逻辑一致性(指地理数据关系上,包括数据结构和拓扑性质等方面的可靠性)、数据完整性和表达形式的合理性。

影响空间数据质量通常有下述几个因素:

其一,地理空间现象自身存在的不稳定性。例如,某种土壤类型边界划分有模糊性、某种土地利用类型边界变动频繁等。

其二,人类认知、测量和表达水平的限制导致的误差。例如,在一些学科中许多概念还没有取得共识,这种不一致性必然导致数据测量误差的产生;数据

测量仪器的精度或误差(如遥感器的分辨率、几何畸变和辐射误差,以及GPS等测量仪器的精度或误差等),以及测量的操作误差和偶然误差,都会带来质量问题,地理空间实体的图形表达的合理性程度,也与精度有关;地图制作本身会带有的误差,包括控制点误差、投影误差等;纸质地图变形也会导致误差等。例如,温度不变,若湿度由0%增至25%,则纸的尺寸可能改变1.6%;而且,纸的膨胀率和收缩率并不相同,回复到同样的温湿条件,图纸不能恢复原有的尺寸。

其三,空间数据处理中的误差。地图坐标变换、投影变换、矢量格式和栅格格式的数据转换等都可能出现误差;制图概括时的个人随意性和不同制图概括人员之间的差异,会导致结果的明显差别;拓扑编辑时所要求的咬合、连接等操作会使图形变形;不同数据层(或图层)之间的匹配会导致不吻合数据的修改;进一步的图层叠加运算会产生空间位置和属性值的变动;不同来源、类型的各种数据集的相互操作过程会产生误差;地图等可视化表达过程中为适应视觉效果,需对空间对象的符号和注记等进行调整等。误差还可以传递和扩散,数据处理的各个过程中,前一过程的累积误差可能成为下一阶段新的误差起源。

地图数字化和扫描后的矢量化处理数字化过程会产生一系列误差。例如,采点的位置精度和采点方式(流模式或点模式)不同会影响结果;不同的数字化要素对象也会导致误差,如粗线比细线更易引起误差,复杂曲线比平直线更易引起误差等;数字化操作人员的技术与经验不同,也会导致较大的差异;数字化仪和扫描仪本身的精度会带来影响;扫描数字化先扫描,再矢量化也会产生误差等。

其四,空间数据使用中的误差。主要包括两个方面:一方面,不同用户对同一种空间数据内容的解释和理解可能不同。例如,对于土壤数据,城市开发部门、农业部门和环境部门对某一级别土壤类型的内涵的理解或解释会有明显差异,从而导致相应的文档说明或元数据的差异或不确定性。另一方面,不少地方缺乏元数据工作、缺少文档,诸如缺少投影类型、数据定义等描述信息,也往往导致用户对数据的随意性使用或误差扩散。

由于如此多的原因,地理空间数据一般都存在大量的质量和精度问题。

空间数据质量和精度方面的数据处理工作,主要有三个方面:

第一,重视数据质量的跟踪和总结工作,计算出各项精度指标数据并作好记录,建立或补充元数据文档,包括记录数据处理过程中质量的变化。

第二,进行空间数据质量的评价,即利用空间数据质量标准,从各种要素,对数据所描述的空间、专题和时间特征进行评价。

第三,更重要的是,在每一项GIS系统项目和工程中严格把握质量,进行空间数据质量控制。空间数据质量控制常见的方法有:传统的手工方法,即将数

字化数据与数据源进行比较 图形部分的检查可采用目视方法，或绘制到透明图上与原图叠加比较；属性部分的检查可采用与原属性逐个对比或其他比较方法等

进行空间数据质量控制的另一个有效方法是元数据方法，即通过数据集的元数据中所包涵了大量有关数据质量的信息，检查数据质量，跟踪元数据了解数据质量的状况和变化。

还有一个有效方法是利用空间数据的地理特征要素自身的相关性来分析数据的质量 例如，山区河流应位于低谷，或位于等高线的外凸连线上等。有经验的地理学家可以从非熟练人员绘制地图中发现很多的问题，就是这个道理。为此，可以建立一个有关地理特征要素相关关系的知识库，以备各空间数据层之间地理特征要素的相关分析之用。

数据质量控制是个复杂的过程，应从数据质量产生和扩散的所有过程和环节入手，分别用一定的方法减少误差。总之，数据质量控制应体现在数据生产和处理的各个环节。

3. 其他 GIS 特色的数据处理

这里，其他 GIS 特色的数据处理指表 4－1 第四大类“GIS 特色的数据处理”中的最下面的五个栏目。它们大部分都是今后要用专门的章节来讲授的内容。

首先，“地理坐标转换和地图投影”，包括地理坐标系、其他坐标向地理坐标转换和不同地理坐标之间的转换、地表曲面空间向平面上投影、精校正和坐标配准等内容，将在下一节，即本章第二节讲授。

“空间数据内插”，指通过已知点或分区的地理空间样点数据，推求任意点或分区数据的处理及其方法，将是本章第三节的内容。

“三维立体处理”将在本章第四节讨论。如前所述，地理空间数据的三维立体处理或 GIS 的三维模型，具有擅长表达不规则地表形态的特点。

“地图输出之前的预处理”，即地图符号、注记和图例等 GIS 特色制图处理，也是 GIS 特色的数据处理的一部分。这部分内容将在本书第六章讲解。

只有“矢量与栅格数据的转换”是本书已讲过的内容（见 § 2.4.2）。顺便说明，现在不少图形图像软件也具有矢量与栅格数据转换的功能，但是，GIS 中的矢量与栅格数据转换，涉及到有关地理属性数据挂联的操作，仍然是 GIS 特色的内容。

4. 广义的空间数据处理

最后简要说明表 4－1 中的第五大类“广义上的空间数据处理”。这些内容在 GIS 书籍中通常用专门篇幅来讲授，且一般不放在“空间数据处理”的名目下。本书中也一样。下面逐条说明之。

“数据采集和输入”和“数据显示与输出”的内容，本书已在第二章讲解，只

有地图符号、注记和图例等地图绘制与输出的内容，不仅重要，而且丰富，需要进一步详细讨论，另列为第六章。我们在第二章解释"数据采集、输入、显示和输出"，主要原因是从教学上考虑：我们认为，这部分内容应当是学习矢量与栅格数据结构的必要部分。因为学习两种数据结构时，很自然地想一鼓作气地了解这两种数据分别是怎样获取、显示和可视化表达的，这有利于较透彻理解矢量与栅格数据结构的基本概念。

"数据存储、组织与管理"，本书已在第二章（部分）和第三章讲述。这部分内容虽然从广义上说也是一种数据处理，但在任何书籍上都作为专门章节来编写。不过，数据管理内容的章节名称，多数 GIS 书籍采用"地理空间数据库"或类似其他用语。

"空间查询和分析"是本书下一章，即第五章的内容。空间分析是地理信息系统的重点，所有 GIS 书籍都有分量不轻的空间分析篇章。不过应当说明，"空间查询和分析"与（狭义的）"空间数据处理"之间并没有截然界限，一些内容，如三维处理和分析、一部分空间操作和某些查询分析等，既可以看作是空间分析空间数据处理，也可以看作是空间数据处理。

§4.2 地理坐标变换及有关问题

地理坐标系是地理空间和空间对象几何表达的基础。本节讲解 GIS 运作中所涉及的地理空间坐标及其变换的问题，包括人类怎样基于地球形状建立地理坐标，以及不同坐标单位所带来的地图输出比例尺问题。

4.2.1 地球表面形状和地理坐标

1. 地球表面形状的几级近似表达

地球表面形状指地球固、液态自然表面，或海陆自然表面的形状。自然表面的形状十分复杂，必须用相对简单的近似环球曲面作为表达或测量的基础。这样的环球曲面有三个：大地水准面、地球椭球体和地球正球体（表面）。三者对海陆自然表面的近似程度从高到低，构成了地球表面形状的三级近似表达。

大地水准面是平均海面（物理上相当于静止海面），及其在陆地下的延伸所构成的一个闭合的环球水准曲面。作为一个处处同重力垂直的水准面，或地球重力场的一个等位面（其上重力势能处处相等），大地水准面的形状能够通过重力测量或地球重力场中的人造卫星运动来确定；同时，它又是重力测量和高程起算的基准面：地表任一点至大地水准面的铅垂距离，就是该点的绝对高程（俗称海拔高度）。我国目前采用青岛验潮站 1953—1979 年验潮资料所计算的平均海平面为统一的高程起算面，称为 1985 年国家高程基准。

作为平均海面，大地水准面不仅能很好表达占地球自然表面70%以上的海面形状，作为高程基准，它又能成为把握陆地表面形状的基础，因为如果已知大地水准面的形状，加上陆地表面各经纬度地点的海拔高度，就等于掌握了地球自然表面的实际形状。因此，大地水准面是地球表面实际形状一种很好的近似。大地测量学中，如果不特别指出是海陆自然表面的话，"地球的形状"常常就是指大地水准面的形状。

大地水准面又远比自然表面平滑、单纯，易于考察。测量表明，大地水准面形状十分接近于一个旋转椭球面。旋转椭球面是一个椭圆绕其长轴或短轴旋转而形成的封闭曲面，是一个易于把握的较简单规则曲面。为此，人们通过测算，选定一个同大地水准面符合得最理想，或者说最逼近的旋转椭球面，并在其基础上表达大地水准面的形状。这样的旋转椭球面所包围的形体，称为地球椭球体。

地球椭球体如图4-1所示。椭球体短轴 b 同地球自转轴重合，称为地球的极半径；其长轴 a 即地球的赤道半径；地球椭球体上的赤道和所有的纬线都是正圆，所有的经圈都是长轴为 a、短轴为 b 的椭圆。可见，地球椭球体的确是一个易于掌握的较规则形体，其形状只需要二个参数 a 和 b 就能确定。

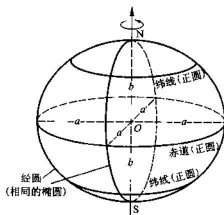


图4-1 地球椭球体

大地水准面与地球椭球体表面的相互逼近，意味着大地水准面在规整的地球椭球体表面附近起伏。大地水准面上任一点至地球椭球体表面的距离 N，称为大地水准面差距，它体现大地水准面上各点对地球椭球体表面的偏离。研究表明，N 值的大小通常不到一、二百米，这个数量对地球来说，实在非常微小。可见地球椭球体表面同大地水准面是非常逼近、相当贴合的。这样，人们如果

测得参数 a 和 b，掌握了规则的地球椭球体表面形状，那么，再加上微小的大地水准面差距 N，就等于把握了大地水准面的形状。

可见，地球椭球体是地球形状的很好的二级近似，人们将它作为测绘部门建立大地坐标系的基础。

在精度要求不高的场合，地球形状还可以进一步简化为正球体。事实上，地球椭球体长短半轴之差仅为 21.4 km，这个差值约为地球半径 6 371 km 的近 1/300。若按真实比例作图，并取 $a=3$ cm，那么，b 应为 2.99 cm。所以，地球椭球体"扁"的程度，远没有图 4-1 所示的那么大，地球椭球体看上去实际是相当圆的。因此，人们以地球正球体表面作为地形状的第三级近似。

地球形状的三级近似如图 4-2 所示。图中，各方框间的箭头及其文字给出各级近似之间的关系；在各方框的下部，给出表示各级近似之间差异大小的量级。

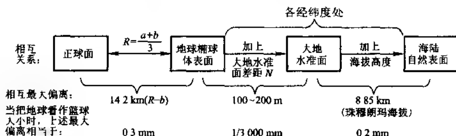


图 4-2　地球形状的几级近似

2. 两种地球椭球体

在上述三层次中，地球椭球体与地理坐标关系最密切，还需进一步说明。如上所述，地球椭球体表面是同大地水准面最逼近、最贴合的旋转椭球面。但是，由于定位方法或"最逼近"的方式不同，地球椭球体分为两类：参考椭球体和平均椭球体。前者只在地球局部范围内贴合大地水准面；后者则在全球范围内贴合大地水准面。

参考椭球体是具体国家"大地测量网"的基础；而大地网点及相应的测量无法在海上和难于涉足的陆地上布设或进行，更不能随意跨越国界。因此，各国通常只能详细把握本国的大地水准面，测算出最贴合本国大地水准面的参考椭球面。所以说，参考椭球体是局部定位的地球椭球体。我国在 20 世纪 50 年代采用从苏联延伸过来的克拉索夫斯基椭球体为参考椭球体；但多年测算发现，这个较适合苏联国情的参考椭球不适合于我国。于是，1978 年我国在西安重新定位，并选用更新的参数，形成了我国 1980 年国家大地坐标系（简称 1980 年西

安原点)。我们在GIS工作中常用的,就是这个大地坐标系。

在全球范围绝对定位的平均椭球体概念则来自于重力测量 原来,重力测量中出于重力场理论和数学上的考虑,要求一个质量、角速度与地球相同,旋转轴和中心分别与地轴和地心重合的地球椭球体的表面为理论根据面,且其重力位(势能)数值与大地水准面重力位相等。可以证明,这样的椭球将在全球范围内与大地水准面最贴近,表达地球整体的平均状态,所以称为平均椭球。

由于不能跨国界测量等原因,早先人们曾缺乏世界范围内平均椭球的可靠概念。人造卫星出来后,通过精密观测在地球重力场运动的人造卫星位置及其变化,可以反算地球重力场,确定包括全球大地水准面在内的各重力等位面,并基于全球大地水准面测算出平均椭球。不仅如此,在卫星观测数据足够多时,还可以进一步计算出观测点在全球重力场或地心坐标系中的精确位置,从而为GPS技术的兴起打下理论基础。

这里自然引起一个问题:为什么各国不放弃参考椭球,而采用显然在理论上更优越的平均椭球作为全球统一的地心坐标系统之基础呢?原因主要有二:其一,平均椭球的各物理参数随着现代测量技术的进步而不断变动,而具体国家大地网及与之相连的参考椭球在实践中却不宜经常改动,因为改动意味着大量测绘或地图数据的更换。其二,出于国家利益的考虑,各国也需要保密自己的大地网、参考椭球和详细测量数据;并同时把握参考椭球与平均椭球间的转换参数,保持在理论上与全球地心坐标系之衔接。

3. 三种地理坐标

由于上述原因,在全球观测体系中,有着三类不同意义的地理坐标。

第一类是地心坐标。这种地理坐标依据地心半径(观测地点与地心的连线,如图4-3中的AO)来定义;地心半径r加上地心经纬度λ,ψ,合成地心坐标(r,λ,ψ),该坐标可以很方便地转换为地心直角坐标(x,y,z)。地心坐标通常与平均椭球相联系,表达观测地点在全球地心系统中的具体位置。GPS观测所直接得到的地理坐标,理论上就是这种坐标。欲求得下述大地坐标,还必须经过换算。

起始大地子午面
参考椭球
陆地
B
A
O
O'
A'
ψ
φ
λ

图4-3 地心坐标和大地坐标

第二种是依据参考椭球建立的大地坐标。大地经纬度依据"垂直线"来定义,垂直线是与参考椭球体表面垂直的法线 除两极和赤道上外,垂直线不通过参考椭球中心,如图4-3中的AO'所示。图中可见,大地纬度φ与地心纬度ψ有较明显差异。注意,参考椭球是仅贴合具体

国家、局部定位的地球椭球体、参考椭球的中心与地心或平均椭球中心并不重合;但这种偏离较小,在图4-3中不可能表现出来。

大地坐标就是GIS工作中常用的、具体国家大地测量的地理坐标。因此,大地坐标对我们特别重要,需多说几句。在大地坐标系中,参照图4-3,设某地点B在参考椭球面上的投影为A,则过A点的垂直线与赤道平面的线面角φ就是当地大地纬度。大地经度则是从起始大地子午面至A点所在大地子午面之间的两面角λ。这里,大地子午面是过参考椭球短轴的半平面。地表任一点至参考椭球表面的垂直距离,称为大地高,例如,图中AB即为B点的大地高。大地经度、大地纬度加上大地高,三者合起来确定实际地点相对参考椭球的位置。

第三种地理坐标为单纯的球面坐标,即只有经度、纬度二维坐标,称为天文经纬度。天文经纬度借重力方向(即铅垂线方向)来定义。铅垂线是大地水准面的垂线。天文经纬度主要用于天文观测中,GIS工作者接触的机会不多。

地表任一点都有自己的铅垂线、垂直线和地心半径。由于地球形状略扁,且不规则,这三条线都相互偏离,但偏离的程度不同。一般而言,铅垂线和垂直线相当接近,即大地经纬度与天文经纬度十分接近;垂直线偏离地心半径相对较大,即大地经纬度与地心纬度数值差异相对明显。当把地球作为正球体来处理时,三线合一,三种经纬度之间没有差异。

4.2.2 地图投影和坐标转换

1. 地图投影

地球太大,人置身其中,难识"庐山真面目",为了可视化地研究地表或区域,必须将所关心的地理事物浓缩到一块足够小的、方便于人观察的表面上。当区域范围不大时,地表可以近似地看作是平面,区域及其事物很自然地浓缩到平面上,因而形成了平面地图。但是,当研究的区域范围很大,地球表面的球形分布不可忽略时,要想保持正确的比例关系,地表应当浓缩到一个球面或椭球面上,就像地球仪那样。然而,地球仪携带不便,加上人类长期观察平面图的习惯,人们仍然大量地采用可卷折的平面地图。这就要求将地表曲面上的事物按一定的数学规则,一一对应地转换到平面上。这种在球面与平面间建立点与点之间对应数学关系的方法,称为地图投影。

具体的地图投影(数学)方法很多,常用的就有数十种。不论用哪种数学方法,球面上地理事物投影到平面上,都不可能处处保持相似性,换言之,地图投影必定产生变形,地图上地物的长度、角度、面积和形状必定会部分地与对应的实际地物发生差异。按投影变形性质的不同,地图投影可分为三类:第一,等角投影。此类投影中,投影面与地球椭球表面上相应的夹角相等,使地物形状在图上得以较好保持,但地物长度、面积因地而异的变动较大。第二,等积投影。

投影面上的图形面积与地球椭球表面上相应图形的面积相等,但图形的轮廓形状会产生较大的变化。第三,任意投影。任意投影中,角度、面积和长度三种变形同时存在。其中,有一种"等距投影"能保持一个方向上的等距性质。上述三类投影各有其适用性。

地图投影的几何面可以是平面,或可展开成平面的曲面,如圆柱面和圆锥面。用以投影的几何面称为投影几何面或辅助几何面。按投影几何面的不同,地图投影也可以划分为三个主要类别:方位投影、圆柱投影和圆锥投影,它们分别采用平面、圆柱面和圆锥面作为投影几何面。不论哪种,投影几何面皆可能采取多种与地球的几何配置方式。例如,投影面可与地球相切或相割,分别称为切投影或割投影;投影几何面可与地球自转轴垂直、平行和斜交,分别称为正轴、横轴和斜轴方式。图4-4较好地表现了这三种投影几何面的分类,以及每一类中投影面与地球的多种不同的几何配置的形态。图中PP表示地球自转轴,EE表示地球赤道。

图中可见,方位投影直接采用平面作为投影面,直接将球面上经纬网投影到平面上。不难想象,当采取正轴方式(切或割)时,纬线呈同心圆,经线为自圆心辐射的直线。方位投影在切点或割线上无任何变形,离切点或割线愈远,变形愈大。变形性质可以是等角、等积和等距的。方位投影一般适用于编制具有圆形轮廓的地区,如极地和半球的地图。

圆锥投影以圆锥体面为投影面,先将球面上的经纬网投影到圆锥体面上,再将圆锥体表面展开成一个扇形平面,制作地图时取其中所需部分。当采取正轴方式时,纬线亦呈同心圆,经线为自圆心辐射的直线。在切线或线上无任何变形,离切线或割线愈远,变形愈大;变形性质也可以是等角、等积和等距的。圆锥投影适用于编制中纬度地带沿纬线方向伸展地区的地图。

圆柱投影以圆柱体面为投影面,先将球面上的经纬网投影到圆柱体面上,再将圆柱体面展开成一个矩形平面。当采取正轴方式(切或割)时,经线呈等间距的平行直线,纬线为垂直于经线的另一组非等间距的平行直线。其横轴方式下文还将涉及。圆柱投影中,切线或割线上无任何变形,离切线或割线愈远,则变形愈大;变形性质也可以是等角、等积和等距的。圆柱投影一般适用于编制赤道附近地区的地图和世界地图。

2. 高斯-克吕格坐标

地图学家利用各种投影绘制各种各样的地图。但是,大多数GIS工作绘制和输出的地图一般涉及的投影种类却很有限。这是因为,地球曲率不可忽略的大范围(小比例尺)地图的绘制,常涉及多种投影;GIS软件中虽然有很多种地图投影功能,也带有不少的大范围电子地图,但一般用以学习示范。大多数GIS工作者面向具体的、范围有限的区域,常常仅进行平面地图之间的转换。

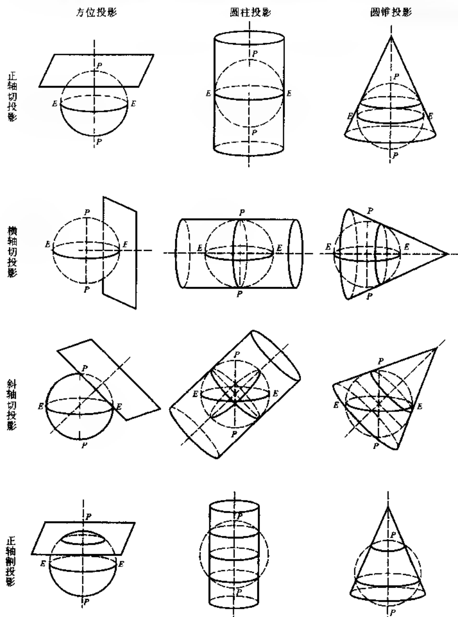


图4-4　按投影几何面的分类及投影面与地球的几种配置
（取自马永立《地图学教程》）

GIS 工作者涉及地图投影主要有两种情形。一种是海上 GIS 项目,需要应用海图。海图一般采用等角正切圆柱投影(这里"正"指正轴),又称为墨卡托投影,其投影方式如图 4－4 上排中间的小图所示,地心为投影中心。不难想象,在这种投影下,经线是均匀的,纬线间距在赤道上最小,随纬度增高而迅速增大。墨卡托投影是等角的,其地图上起点和终点间的直线,能指示通向目的地的等角航线。

不过,海上项目毕竟有限。大多数 GIS 工作涉及地图投影,是因为陆地项目中需要应用国家地形图,因而需要涉及地形图所采用的地图投影方法。在我国和其他很多国家,1:100 万比例尺的地形图采用等角正割圆锥投影,其投影方式如图 4－4 最右下小图所示;而 1:50 万至 1:5 000 比例尺的各级地形图均采用高斯－克吕格投影。由于大多数 GIS 项目涉及比例尺在 1:100 万以下,本小节将重点介绍高斯－克吕格投影。

高斯－克吕格投影简称高斯投影,采用等角横轴切椭圆柱投影,如图 4－4 第二排中图所示。但是,它不是将整个地球一次投影,而是采用分带投影的方式,将地球椭球体表面按地理经度每 3°或 6°分带(每一带像一块"西瓜皮"),并将每一带投影到切于该带中央经线的椭圆柱面上,再将圆柱面展开成平面。每一带投影结果为橄榄形,各自建立坐标网,如图 4－5 所示。图中可见,高斯投影的经纬网形状的特点是:除中央经线和赤道投影为互相垂直的直线外,其余经线的投影为对称凹向中央经线的曲线,纬线的投影为对称凸向赤道的曲线;整个图形呈东、西对称,南、北对称,经纬线均正交。

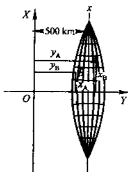


图 4－5　高斯投影结果及坐标

图 4－6 表示国际上共同约定的分带规则,图中,上面若干半橄榄形表示 6°分带,下面若干半橄榄形表示 3°分带。图中可见,6°分带投影从零子午线起,由西向东,每 6°为一带,全球共分 60 带,用阿拉伯数字 1 至 60 标记之。我国从 13 至 23 带,共跨 11 个 6°分带。3°分带投影从东经 1°30′起,也是由西向东,每 3°为一带,全球共分 120 带,用阿拉伯数字 1 至 120 标记之。

注意,图 4－5 和图 4－6 为清晰表达内容,较大地夸大了橄榄形的宽度。实际上,且不谈 3°分带,即使是 6° 分带,也是非常狭长的。这样狭长的"西瓜皮"投影到紧贴着它的椭圆柱面上,变形是比较小的;只有赤道上离中央经线较远的地区变形相对偏大一点。为此,对于制图精度要求高的较大比例尺地形图,要求采用 3°分带。我国规定,1:50 万至 1:2.5 万地形图采用 6° 分带投影,1:1万及更大比例尺地形图采用 3° 分带投影。在这样的规定下,高斯投影的变

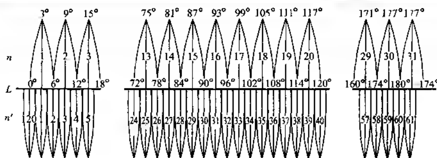


图4-6 高斯 克吕格投影分带

形误差是足够小的，甚至没有超出绘图和量图作业所产生的误差范围。高斯投影还具有公式简单、全球规则化等优点。例如，3°带中有半数中央经线与6°带的中央经线重合，以致3°带计算可直接用于6°带；带与带之间具有同一性，每个带内又有上下、左右的对称性，以致1/4个带的计算可以推广到全球每个投影带。所以，高斯－克吕格投影是一种制作较大比例尺地图的较理想投影方案，适用于广大测区的大地测量，因而为许多国家所采用。

基于高斯－克吕格投影建立的高斯－克吕格平面直角坐标系，如图4－5所示。该坐标系以"米"为单位，并分别以中央经线和赤道投影而成的两直线（互相垂直）为坐标轴，其中，横坐标轴 Y 为赤道投影后的直线，纵坐标轴 X 按我国规定，为中央经线投影后的直线向西平移500km。后者是因为横坐标 Y 向东为正，向西平移500 km能使全部横坐标值均为正值，方便处理。这样，中央经线的 y 值为500 km，即500 000 m。纵坐标 X 向北为正，向南为负；我国位于北半球，x 值全为正值。由于每1度纬度所对应的经线平均为111km或111 000 m，从赤道起到我国30纬度的地区，x 值约为3 330 000左右。可见，高斯－克吕格坐标 x 和 y 的数值，通常分别为以米为单位的7位数和6位数。

高斯－克吕格坐标计算在各带通用，导致不同带间坐标值可能混淆。例如，我国跨11个6°分带，就可能有11个地点有相同的高斯－克吕格坐标。为了加以区别，规定可在 y 值前面冠上所在带号。例如，20带内某点原为642 683 7(m)的 y 值，可化为 $y=20\ 642\ 683.7(m)$。这样就转化为全球通用坐标。

高斯 克吕格直角坐标网格称为方里网。我国规定在各级地形图上均标绘方里网。由于各级地形图依据经纬度统一分幅，在我国地形图的图廓上就有两种坐标网标记，即不仅有经纬网，而且有方里网标记。

高斯－克吕格投影的主要缺点是分带投影所导致的不同带间的相邻图幅

不能直接拼接，致使跨带的区域地图绘制较为麻烦。例如，120°经线穿过浙江省中部，使该省分居20带和21带，经常遇到跨带制图的问题。这时，不得不选取其中一个带，并在一定范围内将该带坐标网延伸到邻带图幅中去。GIS工作中常用基于高斯－克吕格投影公式编制程序来实施这种延伸。在国家地形图上，为了方便邻带间相邻图幅的拼接，规定处于带边缘的地形图应同时标注邻带相邻图幅的方里网。我国1971年版《1:2.5万～1:10万地形图图式》规定，西带的方里网要延伸至东带达30′宽的范围，东带的方里网要延伸至西带达15′宽的范围。

关于地形图的分幅规则和邻带相邻图幅拼接的具体内容，请参阅地图学书籍。

3. 平面地图直角坐标间的转换：坐标校正

大多数GIS工作面向具体的、范围有限的区域，经常涉及平面地图或图像的直角坐标间的转换。这种转换就是要寻求一个关系式

$$x' = f(x,y) \quad 和 \quad y' = f_2(x,y) \tag{4-1}$$

将地图上各点的原坐标(x,y)，转换成新的坐标(x', y')。这个工作我们统称为校正或坐标校正(rectification)。校正可分为两类：一类是将非地理坐标转换为地理坐标，英文常称为georeferencing，这里暂时译为地理坐标化。一个典型的例子是：通过手扶跟踪数字化仪或扫描仪矢量化输入地图，输入设备会自动赋予反映地图实际尺寸的普通直角坐标，通常以英寸为默认单位，一幅地形图输入后的坐标，其数值多在20以内(1inch = 2.54 cm)。例如，数字化的地形图中部某点的坐标可能为$x = 11.3, y = 9.8$等。这种电子地图在需要进行地理分析时，其原有的普通坐标将转换为高斯－克吕格坐标。这就是地理坐标化。遥感图像也常需要地理坐标化，那是将体现栅格中行、列位置的图像坐标（参见§2.3.1.2），转换为高斯－克吕格坐标。

另一种类型的校正称为校准或配准(registration)。校准可在任意两个平面坐标系间进行，不一定非要转为地理坐标。校准工作在GIS中也是常见的，因为实践中，常有同区域不同来源的地图或遥感图像不相匹配的情况；或原有坐标系不同，或坐标系种类相同，但地物坐标值不佳，以致两地图或图像叠在一起时，很多地物，如河流道路，甚至连图幅（框）都不能很好吻合。这时，就要设法将其中一幅图的坐标向另一幅图校准。校准的目的是使两图间匹配或套准，而不关心是否注册到地理坐标系中。

地理坐标化和校准这两种坐标校正，虽然在内容上有差异，但二者的技术路线相似。下面就来讨论技术方法。注意，坐标校正与地图投影不同：地图投影也要有一个像式(4－1)那样的转换关系，但是，它是将（椭）球上的坐标(x, y)，转换成平面坐标(x', y')，其关系式基于一定的投影法则，有确定的数学表

达公式;而坐标校正是从平面坐标到另一个平面坐标,其转换关系没有确定的一成不变的数学表达式,需要通过地图的实际情况,即通过地图中可用的地面控制点来实现。

这里,地面控制点(ground control point,简称 GCP,在 Arc/Info 中称为 Tic)不是指测量学中的大地控制点,而是地图或图像校正中,同时已知原坐标数值(x, y)和新坐标数值(x', y')的地面点。例如,以国家标准分幅绘制的地形图和土地利用图等的四个图廓点,或其他特殊点,包括可见标志的大地控制点等,是已知经纬度坐标或高斯-克吕格坐标的点;这些点在地图数字化输入后,又有了输入设备所赋予的普通(英寸)坐标。因此,这些同时具有两种坐标系之坐标值的点,就可以作为地面控制点 GCPs(GCP 的复数)。又例如,在进行两幅电子地图或图像的校准时,可采用两图中的公共标识点,如明显的十字路口、河流交点和其他明显标志物,作为同时已知两种坐标值的地面控制点 GCPs。

地面控制点是求解转换关系的基础。为简单起见,假定式(4-1)为线性关系:

$$x' = a_1x + b_1y + c_1 \quad 和 \quad y' = a_2x + b_2y + c_2 \qquad (4-2)$$

求解这个关系,就是要通过同时已知(x, y)和(x', y')值的地面控制点 GCPs,来求解式中的 6 个系数 a_1、b_1、c_1、a_2、b_2 和 c_2。注意,每一个 GCP 可以提供 2 个已知(x,y)和(x',y')的方程,3 个 GCPs 即可提供 6 个已知(x, y)和(x', y')的方程,从而可解出(4-2)式的 a_1、b_1、c_1、a_2、b_2 和 c_2 6 个系数。这样,图中任意点的(x, y),就可以通过式(4-2),转换为新坐标(x',y')。

在实际工作中,由于各种来源的地图误差普遍存在(参见 §4.1.3.2),应当选取 4 个或更多的 GCPs 来求解(4-2)式。换言之,采用最小二乘法,用 8 个、10 个或更多的方程式,来求解 6 个系数 a_1、b_1、c_1、a_2、b_2 和 c_2。这样求得的,实际上是 6 个系数的最佳值。这方面一个最常见的例子是:在对标准图幅的输入地图进行坐标转换时,采用四个图廓点作为 GCPs,即利用 8 个已知(x, y)和(x', y')的方程,来求解式(4-2)的 6 个系数。

式(4-2)是线性关系。平面地图间的坐标校正也可采用非线性关系。例如,当采用 2 次多项式时,转换关系式(4-1)将为:

$$x' = a_1x^2 + b_1y^2 + c_1xy + d_1x + e_1y + f_1$$

$$y' = a_2x^2 + b_2y^2 + c_2xy + d_2x + e_2y + f_2 \qquad (4-3)$$

式(4-3)需要求解的系数为 12 个。因此,至少需要 6 个 GCPs,以提供 12 个已知(x, y)和(x', y')的方程。再进一步,若转换关系采用 3 次或以上的 n 次多项式,不难证明,所需的最小 GCPs 的数目 k 为

$$k = (n+1)(n+2)/2 \qquad (4-4)$$

例如当 n = 3 和 4 时,所需最小 GCPs 数目分别为 10 和 15 个。不论 n 为多少,

实践中也通常需要多选择一些 GCPs 来求解最佳转换关系,消除部分随机误差。

现在很多软件都有解线性方程组的现成功能,采用多项式拟合已较为方便。最后提一下,在遥感图像校正中,坐标变换后还需要对变换后图像像元的属性值进行某种调整。该项工作称为重采样(resampling,见 §4.3.4.2)。

4.2.3 其他有关问题

1. 关于坐标校正的讨论

上面讲到,在坐标校正中,若采用 n 次多项式作为转换关系式,所需的最小 GCPs 的数目为 $(n+1)(n+2)/2$;但我们却反复强调应选用多于这个数目的 GCPs,通过最小二乘法来求解转换关系。为什么呢?

由最小二乘法决定的转换关系亦即回归曲线。本书不拟介绍最小二乘法原理,仅给出图 4-7 加以形象说明。该图仅以一维 x 坐标来示意,画出从 x 向 x' 坐标转换的 2 次转换曲线或回归曲线。图中给出 15 个 GCPs 样点,很明显,按 15 个样点分布,曲线 1 应是从 x 值转换为 x' 值的最佳转换曲线,15 个样点分布在曲线 1 两侧,意味着曲线 1 消除了部分样点的随机误差。但是,若仅采用其中 6 个 GCPs,例如采用图中 6 个黑点来求解,那么,所得到的转换关系就不是曲线 1,而是转换曲线 2。显然,曲线 2 仅反映 6 个样点的分布,将样本的随机误差拟合了进去,不是比较合理的结果。

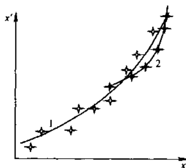


图 4-7 GCPs 与转换关系曲线

再一个问题:在上述坐标校正中,多项式转换关系的次数是否愈高愈好呢?不然。仍参照图 4-7,根据式(4-4),15 个 GCPs 正好足够求取一个四次多项式的坐标转换曲线。这意味着,该四次多项式曲线必定(波动地)恰好通过 15 个 GCPs 的每一个样点,"拟合率"可达 100%。但显然,那样的转换关系反而不可信,因为它把 15 个样点的随机误差也"拟合"进去了。相反,从 15 个样点中间穿过的 2 次多项式曲线,即图 4-7 中曲线 1 反而合理。由此可见,多项式转换关系的次数尽可能取低一些为好。为此,实践中较好的做法是首先取一次多项式,即线性关系来转换,用足够多的 GCPs 来求取转换关系;然后就此进行观察、分析,实在有必要时才"升级",采用二次或以上多项式。

观察、分析转换关系之质量的一个较常用方法,是计算每一个 GCP 样点相对回归线的偏离。偏离一般用均方根表示:

$$均方根 = \sqrt{\Delta x^2 + \Delta y^2}$$

这里，Δx 和 Δy 分别为 GCP 样点在 x 和 y 方向相对回归线的偏离。均方根可用来辅助判断 GCP 样点的质量，如果某个 GCP 样点的均方根特别大，那么，该点可能有问题，宁可不作为样点使用。去掉有问题的 GCPs，用剩下的 GCPs 来求取转换关系，可以提高转换关系的质量。如果有较多 GCPs 的均方根偏大，那就可以考虑将多项式升级，例如不用线性关系而采取二次多项式。

现在一般的 GIS 软件和遥感图像处理软件都有现成的坐标转换或校正的功能，让用户选取转换关系及其次数，并给出转换后总均方根和每个 GCP 的均方根等参数，用户可以借以方便地剔除有问题的地面控制点，并再次进行尝试，非常方便。但是，软件功能远不能代替用户自己的误差分析或精度分析。事实上，误差分析或精度分析还包含很多深层次内容，研究空间还相当大，这里就不进一步讨论了。

2. 引进国外 GIS 软硬件带来的有关问题

由于我们至今仍然较多地使用进口的，特别是美国的软硬件，有时会遇到进口设备的地理坐标功能和所带的示范地图不符合中国国情的情形，这里也举几例说明。首先，如上所述，美国等国家的 GIS 软硬件的默认度量单位，常设为英制，如长度采用英寸、英里等。当然，用户可以变更设置，例如设为 SI 制，如厘米单位等。但是，作者认为，任何设置都不是一劳永逸的，在转换为高斯－克吕格坐标以前，带着英制单位运行未尝不可，关键在于自己心中有数。此其一。

其二，GIS 和遥感图像处理软件一般都给出众多的地图投影和坐标系统供选择，但是，软件生产商不一定会注意到别国的国情，使我们有时应用不便。例如，在一些美国软件中找不到高斯－克吕格坐标或投影，因为它们将这种投影称为 UTM，即通用的横切墨卡托投影。UTM 是美国所用的两种大地平面坐标之一。用户只有知道这一点，才可能利用软件的 UTM 功能进行普通坐标向高斯－克吕格坐标的转换。又例如，从我国卫星接收站购买的，已注册高斯－克吕格坐标的 TM 遥感图像数据，其参考椭球体注明为克拉索夫斯基椭球体；而一些美国软件不认识这种椭球体。当然，只要不涉及有关投影，软件照样运作。

此外，进口 GIS 和遥感软件通常带有不少的示范地图；但这些地图也是面向他们自己的国家。例如，某些进口 GIS 软件所带的全景世界地图以 0°经度线（过英国的本初子午线）为中央经线，这种地图上，欧洲处于较中间位置，美洲稍偏，而中国处于变形很大的边缘位置。这样的世界地图就不适合我国。我国地图出版社至今采用 1963 年设计的世界地图投影，以东经 150°为中央经线，使我国位于较中间位置，并保持太平洋的完整。又如，有的 GIS 软件所带的、不同国家赋不同颜色的世界地图，将中国大陆和台湾省赋两种颜色，这种地图不符合

个中国的原则，应注意。

3. 计算机内的比例尺与坐标单位

还有一个有 GIS 实用意义，但一般 GIS 书籍未能提及的问题，是计算机内的比例尺与坐标单位问题。众所周知，计算机内的地图是可以随意缩放的，且 GIS 软件都具有屏幕显示自动适合地图范围的功能，以致地图经过坐标变换，显示出来与从前没有任何差别。例如，一幅坐标数值多在 20 以内（以英寸为单位）的数字化地形图，与转换成高斯－克吕格坐标，数值为 6～7 位数（以米为单位）的同一幅地形图，在计算机中显示完全一样，因为对前者，GIS 软件让屏幕自动显示 20 左右数值的范围；而对后者，屏幕自动显示一个 6～7 位数表示的范围。

由此缘故，GIS 中的比例尺具有双重含义。一种是输入前（纸张）地图的实际比例尺；另一种是地图在计算机里的比例尺。作者认为，GIS 中涉及计算机内比例尺问题的情形主要有两种：屏幕显示和地图输出。屏幕显示的比例尺可以定义为屏幕尺度与地图所表示的实际区域尺度的比值。例如，全屏显示 1∶1 万县域地图的显示比例尺，就小于全屏显示 1∶1 万乡域地图的显示比例尺。前者同后者相比，将显得较为密麻，即地图负载量或信息量较大。现在不少 GIS 软件将实际比例尺不同的地图分为不同的图层存放，并具有在屏幕上缩放地图时自动增减图层的功能，以保持各种屏幕显示时适当的图面负载量。屏幕显示比例尺及其与屏幕上地图图面负载量的关系，是传统地图学所没有的课题。

地图输出比例尺指 GIS 软件在输出时，要求用户指定的比例尺。这个比例尺通常不等于地图的实际比例尺，它取决于三个要素：原图的实际比例尺、计算机中地图坐标及其单位，以及输出设备的默认单位。后面两个要素起作用是因为：地图输入后会进行坐标变换，变换前后的地图在计算机中同样显示，但当涉及到地图输出等实际应用的几何计算时，不同坐标及其单位就会导致差异；而计算机只管运作数字，它不明白坐标及其单位的意义，也不明白坐标单位与输出设备的默认单位之差异和关系。因此，为了正确地运作，用户必须进行人工干预。让我们循序渐进地举例说明。

例 1： 一幅 1∶10 万比例尺的地形图输入到计算机中，以英寸为单位，将这幅电子地图用绘图仪（默认单位也为英寸）输出，应指定多大的输出比例尺，所绘制的地图仍为 1∶10 万比例尺？

答案：输出比例尺应为 1∶1。

分析：该图输入计算机后，坐标值的“1”代表 1 英寸，坐标数值范围（例如约为 20 英寸）反映该图实际大小；而绘图仪单位为英寸，意味着绘图仪也把计算机里的“1”作为 1 英寸对待。因此，1∶1 的输出比例尺，意味着原图原封不动地输出，自然还是原来的大小和原来的实际比例尺。

例 2:其他条件同上例,但绘图仪的默认单位被设置成厘米,这时应指定多大的输出比例尺,所绘制的地图仍为 1:10 万比例尺?

答案:输出比例尺应为 2.54:1。

分析:计算机里的地图单位仍为英寸,即数据"1"代表 1 英寸,但这时绘图仪把数据"1"作为 1 厘米对待,即把原为 1 英寸的尺度绘成 1 厘米大小,使实际输出地图的尺寸缩小到 1/2.54。因此,同上例比较,为了原封不动地输出原图,必须将 1:1 的输出比例尺提高为 2.54:1。

例 3:其他条件同例 2,但地图坐标已转为高斯–克吕格坐标,这时应指定多大的输出比例尺,所绘制的地图仍为 1:10 万比例尺?

答案:输出比例尺应为 1:1 000

分析:地图坐标已转变为高斯–克吕格坐标,意味着电子地图的 6~7 位数的坐标值已表现地球上实际区域的大小;且其单位为米,即数据"1"代表 1 m。如果这时绘图仪的输出单位也是米,那么,就应当指定实际比例尺 1:10 万为输出比例尺。但是,这时绘图仪的默认单位为 1 厘米,即把坐标值"1"(米)作为 1 厘米绘出,使绘出的地图缩小到 1/100。因此,输出比例尺应在 1:10 万的基础上增大 100 倍,即为 1:1 000。

例 4:其他条件同例 3,但绘图仪默认单位为英寸,这时应指定多大的输出比例尺,所绘制的地图仍为 1:10 万比例尺?

答案:输出比例尺应为 1:2 540。

分析:其他情况同例 3,但这时绘图仪的默认单位为 1 英寸,即把坐标值"1"(米)作为 1 英寸绘出,使绘出的地图缩小到 2.54/100。因此,输出比例尺应为 1:10 万 ×100/2.54 = 1:2 540。

最后再给出另一种类型的,但有联系的实例作为课后的进一步思考题。该题目是作者在若干年前,在 1:10 万土地利用现状图缩编为 1:25 万图的过程中,进行计算机制图综合(概括)研究时所处理过的实际问题。该问题的条件是,1:10 万土地利用图输入计算机,采取数字化仪所赋予的英寸单位,制图综合后要输出为实际比例尺 1:25 万的地图(不论绘图仪单位如何)。问:该数字化的 1:10 万土地利用图上面积为多大的图斑,在将来输出的 1:25 万地图上面积等于 2 mm^2?

§4.3 地理空间数据插值

4.3.1 基本概念

1. 空间数据插值的概念和必要性

空间数据插值指通过已知点或分区的数据，推求任意点或分区数据的处理及其方法，它是地球科学和相关领域中，具有基础意义并广泛应用的一种空间数据处理类别。数据插值的运作对象，通常是二维地理分布的某种地学或相关学科领域的空间变量或属性值。设此变量为 z，按某种关系随地理位置 (x,y) 在地理空间分布，且区域内共有 n 个已知数据，则空间数据插值就是要基于已知的

$$z(x_i,y_i) \qquad i=1,2,3,\ldots,n$$

求取该区域空间变量 z 分布的某种关系式：

$$z=f(x,y)$$

从而可推求区域内其他地点 (x,y) 处该空间变量 $z(x,y)$ 的值。一般而言，数据插值要以数据质量的损失为代价，因为所推求的分布关系不可能完美地模拟或表达空间变量 z 分布的真实规律。因此，数据插值常常是一种不得已而为之数据处理，绝非多用为好。

空间数据插值之所以有非常广泛的应用需求，是因为在地球科学各分支和很多相关领域，由于种种原因或条件限制，数据不仅在精度或质量上有很多问题(见§4.1 3.2)，而且在空间分布上也常常难尽人意。后者正是导致这些领域需要进行数据插值的根源。

第一，由于自然环境较恶劣地区难以采集数据，或观测仪器(如遥感器)故障导致部分数据缺失等原因，导致数据分布不完整，部分区域数据缺失，需要进行空间数据插值。

其次，数据虽然在整个地区都有分布，但也可能分布不均，部分地区样本过于稀疏，或者整个区域的数据都比较稀疏，也需要进行插值以加密数据。

第三，并不一定因为条件限制，而是由于对空间变量分布规律不够了解，人为的数据采样方式不适合该分布规律，例如在分布有跃变的地区采用均匀采样方式等。

第四，即使是看来不错的数据，当需要转换数据格式，进行坐标校正，或在应用对象变化等情况下，也可能需要进行数据插值。例如，遥感影像数据进行坐标校正或分辨率调整，TIN 数据转为栅格形式(或相反)等，都需要进行数据

插值。

空间数据插值包括空间数据内插和数据外推两种类别。前者通过已知点的数据，推求同一区域其他未知点数据；后者则是通过已知区域的数据，推求其他区域数据的方法。一般而言，空间数据内插应用得较为广泛，空间数据外推的范围或能力有限。本节下面将要介绍的插值方法，主要是面向空间数据内插的，其中一部分也适用于空间数据外推。

2. 空间变量分布与不同空间分布的数据源

为了减少空间数据插值带来的数据质量损失，空间变量数据源的分布最好能比较适应空间变量的分布特点。众所周知，空间变量的分布或者采取连续变化的方式，包括缓变与剧变的连续性变化，或者采取跃变方式（一个区域内可能兼有这两种情形）。为了数据源的分布能适应空间变量分布的变化方式，一般而言，在属性连续渐变、缓变型区域采样点可以少一些，均匀一些；在变动频繁或剧烈的区域，采样点应当密一些；在属性分布有跃变的地区应当沿跃变带密集采样等。有一些学科在采样时就注意到这些要求，例如，根据需要采用规则采样、随机采样或二者结合的方法；河流、山坡剖面的测量采用规则断面采样的方式等。当然，现实中的空间数据采样由于条件限制不一定能达到上述要求。

GIS 实践中常见的数据源在空间分布上有下述几种类型：

第一，栅格数据，包括卫星遥感和航测的卫星影像和航测数据，以及卫片和航片的扫描数据，是一种密集的规则分布数据。其特点是，在像元级微观尺度上，这种数据是规则分布于矩形格网上的离散数据；但由于其密集性，数据密密麻麻地覆盖整个区域，以致从区域宏观尺度上，可以看作是连续分布的数据。这种数据的主要优点是详细、规则，便于计算处理，但不能很好适应空间变量分布的差异：在空间变量渐变、缓变型区域，数据常有冗余；在变动频繁或剧烈的区域，可能还不够密集。

第二，较均匀的采样数据。很多学科在没有明显跃变地区进行野外测量数据采样时，采用较均匀的采样方式。采样的数据虽然不像栅格数据那样规则，但它们是较均匀的不规则数据，能通过数据插值，以较小的质量损失代价，转化为规则的栅格数据。

第三，不均匀或不规则数据，包括野外测量的不规则采样数据，和手扶跟踪数字化仪或扫描矢量化产生的地图数据。在数字化地图数据的情况下，空间变量就是几何点、线输入后所赋予的属性值（参见§2.2.4）。例如，用数字化仪手扶跟踪等高线生成 TIN 数据，实际上就是沿等高线采集的，具有随机分布特征的不规则 x，y 坐标样点，为等高线所输入的高程属性值，就是 TIN 数据的空间变量（参见§2.4.3.1）。因为数据样点主要是沿线（线和多边形边界）分布，在地图上有线条处，样点密集，在无线处一片片空白；再加上作业员本身沿线处

理时的随机性，导致数字化的 TIN 数据的不规则分布

数据源分布的不规则性是否为缺憾，不能一概而论，应具体情况具体分析 有些不规则分布具有适应空间变量真实分布之差异的优点：在渐变、缓变区数据较少，在剧变、跃变区样点较密集 例如，数字化采集的 TIN 数据就常有这种特点，等高线较密、较弯曲处通常正是数字化采集点较密集的区域。

尽管如此，不规则数据在计算机处理上较麻烦，规则栅格数据的详细、规则和便于计算处理的特点有时仍然是起决定作用的 因此，在空间数据插值处理中，将不规则数据内插到规则网点上的插值处理，是一项常见的工作，尽管这种插值处理要以数据质量损失为代价。GIS 中常见的空间数据插值工作还有：将规则或不规则的样点数据内插到等间隔数值（常为整数），以形成等值线；或进行加密内插或网格扭变所导致的内插等。

3. 空间数据插值基本思想和理论假设

空间数据插值方法可以从不同的角度分型。首先，从基本思想或理论假设上，空间数据插值主要有四种类型：

第一，基于专业知识基础进行空间插值。例如，根据一个或多个空间变量的经验方程进行整体空间插值，本节后面要讲的变换函数插值，就属这种类型。基于专业知识的方法特别适宜于数据样本不够多的情形，但前提是对该空间变化的自然过程或现象的机理有所理解，能辅助进行空间插值。

第二，用一个数学关系（方程或某种数学函数的级数）来模拟已知样点上空间变量的分布。与第一种方法不同，这种方法（及下面两类方法）没有能辅助空间插值的专业知识基础，因而往往不得不对空间变量的分布及变化作一些假设。本类方法假定所处理的区域变量 $z(x,y)$ 是连续变化的，并认定：任何一个连续的数学表面总是可以用一系列有规则的数学表面的总和，即二元函数之级数，以任意的精度进行逼近。本节后面要讲的趋势面方法和局部函数逼近法等，即属此类型。

第三，邻近样点数值决定法，通常适宜于局部的插值处理。这种方法也可说是按距离插值，即利用待插点邻近的样点的数值来决定待插点的数值，并且在多数情况下，愈近的样点数值对插值结果的贡献愈大。这种方法也假定空间变量是连续的，并假定：空间位置上愈靠近的点，愈可能具有相似的特征，空间变量的值愈可能相近；而距离愈远的点，其特征相似的可能性愈小，空间变量值的差异应当愈大。本节后面要讲的一部分局部内插法，属于这种类型。

第四，边界插值法。这种插值方法基于学科分类原则，在同类或同质区域，只认定一个代表分类意义的空间变量值，忽略同类区域内部的差异，即认为变化仅发生在边界上。这种插值方法主要用于地质、土壤、植被或土地利用等领域的等值区域图或专题地图的处理。

上述四种基本思想或理论假设是本书特别归纳整理的。不少 GIS 书籍也论述空间插值的理论假设，但并未像上面这样分类。

4. 整体插值法和局部插值法

大多数 GIS 书籍都从插值范围的角度，将空间数据插值方法区分为整体插值法和局部插值法。整体插值方法用研究区所有采样点的数据进行全区特征的拟合；局部插值方法仅仅用邻近的数据点来估计未知点的值。整体插值方法有时不直接用于空间插值，而是用来检测空间变量分布的总趋势及其偏离部分，并将后者，即短尺度的、局部的变化看作随机的和非结构的噪声，从而可能丢失一部分有用信息。局部插值方法能弥补整体插值方法的缺陷，可处理局部较细微的变化，而不受插值区域其他点的插值影响。

本书下面内容也按一般 GIS 书籍习惯展开，分别讲解各种整体插值法和各种局部插值法。但请读者参照上面的四种理论假设分类，注意整体插值法和局部插值法中每一种具体的插值方法所基于的理论假设或基本思想。一般而言，整体插值可以采用上述第一、二和第四种假设；局部插值一般基于上述第三和第二类假设。

4.3.2 整体插值法

1. 变换函数插值

变换函数插值方法根据待插值变量与一个或多个空间变量的经验关系来进行空间插值，是一种常用的整体插值方法。下面以一个研究实例进行说明。

在决定冲积平原的土壤重金属污染的几个主要因子中，最重要是距污染源（河流）的距离和高程两个因子。一般而言，携带重金属的粗泥沙沉积在河滩上，携带重金属的细泥沙沉淀在洪水期容易被淹没的低洼地带；在那些高程较高、离河流较远的地方，洪水频率较低，携带重金属污染泥沙颗粒较少，受到污染较轻。高程和至河流的距离两个量都是比较容易得到的空间变量。因此，可能建立某种重金属含量 z 与高程 h 和至河流的距离 d 之间的经验关系。该关系是一个线性回归方程：

$$z(x,y)=a_0+a_1h(x,y)+a_2d(x,y)+\varepsilon$$

式中，$z(x,y)$ 为某种重金属含量；

a_0、a_1 和 a_2 为回归系数；

ε 为独立于 (x,y) 的正态分布残差或噪声（此项不列出亦可）。上述关系可以就各种重金属含量分别建立，利用这些关系，可以对各种重金属含量进行空间插值，以改进对重金属污染的预测。

待插值的变量与一个或多个空间变量的回归关系通常称为转换函数，它可以应用于很多其他空间变量的数据插值。但应该注意，所有的回归转换函数

都属于近似的空间插值；同时，在建立回归模型时应清楚其物理意义。

2. 趋势面分析和其他数学模拟

趋势面分析假定所处理的空间变量是一个随地理位置(x,y)连续变化的二元函数

$$z = f(x,y)$$

从而，它基于整个区域各样点上的已知数据，寻求一个数学关系（方程或某种数学函数的级数）来模拟该空间变量在整个区域的分布，其方法就像§4.1.2.3中基于GCPs的已知数据寻求坐标转换关系那样；确定模拟样点分布的数学关系后，再根据该关系推求区域内其他点上的z数值。

为什么称为趋势面呢？一方面，从二元函数论可知，在二维空间连续变化的空间变量在几何上可以看作是一个平滑的曲面；另一方面，趋势面分析所模拟出的数学关系是一个回归关系，是一个"平均"地穿过诸样点的，反映样点变化趋势的回归曲面。这个道理也如图4－7所示，趋势面分析不是用恰好数目的样点或求解方程来100%地拟合所有的样点数据，而是用多于必要数目的样点或求解方程，通过最小二乘法，来求解回归关系。

趋势面分析常采用多项式回归方法。多项式回归的基本思想是用二元多项式作为模拟的数学关系$z = f(x,y)$，并用多于必要数目的样点或求解方程，通过最小二乘法来求解回归结果。二元多项式可以是一次的：

$$z = a_0 + a_1x + a_2y$$

也可以是二次的：

$$z = a_0 + a_1x + a_2y + a_3xy + a_4x^2 + a_5y^2 \tag{4-5}$$

还可以是三次的：

$$z = a_0 + a_1x + a_2y + a_3x^2 + a_4xy + a_5y^2 + a_6x^3 + a_7x^2y + a_8xy^2 + a_9y^3$$

一般地，n次多项式为

$$z = \sum_{i+j\le n}(a_n x^i y^j)$$

求解上述多项式关系，就是要求确定方程式中的各系数。以二次多项式式(4－5)为例，该式中待定系数为6个a_0、a_1、a_2、a_3、a_4和a_5，如果我们有6个样点数据，即有6个同时已知x、y和z的样点，就有6个方程

$$z_1 = a_0 + a_1x_1 + a_2y_1 + a_3x_1y_1 + a_4x_1^2 + a_5y_1^2\text{（第1个样点决定的方程）}$$

$$\cdots\cdots \tag{4-6}$$

$$z_6 = a_0 + a_1x_6 + a_2y_6 + a_3x_6y_6 + a_4x_6^2 + a_5y_6^2\text{（第6个样点决定的方程）}$$

通过这6个方程，可解出6个待定系数a_0、a_1、a_2、a_3、a_4和a_5，从而式(4－5)成为可以内插其他点的数值的数学关系。当然，为了求得趋势面，样点数必须大于6个，通过最小二乘法来求解6个待定系数的最佳值，即回归系数。

由此可见，这里的多项式回归与§4 1 2 3所讲解的多项式坐标校正非常类似，求解所需的最少样点数与多项式坐标校正所需的最小GCPs的数目相同，为$(n+1)(n+2)/2$；只是方程的数目减少了一半。这是因为，数据插值只需求解一个空间变量z；而在坐标校正中，需求解两个变量（待定坐标x'和y'），所以解方程的数目多一倍。因此，我们在§4.1.3.1所分析的道理，诸如多项式关系的次数一般取低一些为好、利用均方根来辅助判断样点质量和是否升级多项式等，都适用于这里。

除了趋势面分析外，还有采用其他数学关系（方程或某种数学函数的级数）来模拟已知样点的空间变量分布的整体空间插值方法，如傅立叶级数和小波变换等，后者有时应用于遥感图像处理。但这些方法需要的数据量很大。

3. 边界插值方法

边界插值方法基于学科分类原则，将研究区域划分为若干分区或景观单元，它假设任何重要的变化发生在边界上，边界内的变化是均匀的，同质的，每一个单元内部取同样的空间变量值（图4-8）。这种插值方法主要应用于地质、土壤、植被或土地利用现状等领域的等值区域图或专题地图的处理。

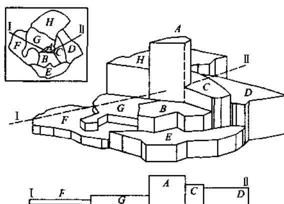


图4-8　边界插值方法示意

泰森多边形或Voronoi多边形（参见图2-16以及§2.4.3 1和§2.4.3.2的小字内容）是一种有特殊意义的、实用的边界插值方法。泰森多边形由任意两个数据样点连线的垂直平分线构成，每一个多边形内只包括一个样点。基于泰森多边形的边界插值方法以样点数值为整个多边形单元的取值。注意到垂直平分线的作用，可知每个泰森多边形内的任意点到自己的样点的距离，始终小于任何到其他多边形样点的距离。因此，泰森多边形插值实质上是一种取最近数据样点之值作为插值的边界插值方法。

GIS 和地学分析中常采用泰森多边形进行边界插值，虽有精度损失，但有快速、简明的优点 事实上，泰森多边形插值的首次应用者是气象学家泰森 众所周知，降水、气压、温度等气象要素本是连续变化的，而用泰森多边形插值的结果，气象要素将呈非连续的跳跃性变化。尽管如此，泰森多边形插值的应用意义仍然得到公认。

4.3.3 采用活动窗口的局部插值法

现在开始分两部分讲局部插值方法 本节先讲解应用较多的活动窗口局部插值；其他局部插值方法合在下面的 § 4 3.4 中介绍 局部插值方法用邻近的样点数据来估计未知点的值。样点数据可以是原来的，也可以在趋势面分析的基础上再生，即在原样点数据中去掉趋势面变化的部分，以所得结果作为局部插值的样点值。

1. 基本方法和技术路线

采用活动窗口的局部插值法又称为移动插值法或逐点内插法，是 ·种常用的局部插值法。该方法选定 系列分布于整个插值区域的、有规律排列的点作为欲进行插值的点（下称待插点），插值时基于每个待插点周围一个邻域内的已知样点数据，依次地确定每 ·个待插点处的数值。在这个过程中，待插点邻域就像 ·个活动窗口随着 个个待插点的处理而移动。因此，这个方法也称为活动窗口方法（moving window），非常类似于遥感图像的局部增强处理中的活动模板或活动窗口。活动窗口法非常适合于将不规则样点数据，包括 TIN 数据，插值到规整的格网点上，转化为栅格数据，因而具有重要的实用意义，有必要较详细加以讨论。

图 4－9 所示，假定我们要将该图所示区域内的不规则样点数据，插值到 ·个正方形格网的每个交点上（网格中心亦可），图中每个小黑点表示样点位置。实际工作中的样点和格网通常比较密集，但图 4－9 为简明起见，仅画出 个 8 × 8 的格网。采用活动窗口法进行插值通常包括下述几个步骤：

第 ·，定义窗口，即插值邻域或搜索范围。定义的方法有多种，其中较简单的窗口是方框或圆框（参见图 4－9，图中画出 ·个圆框和 2×2 的方框）采用这两种窗口编程都比较方便。不论哪种窗口，一般皆以窗口的中心为待插点。

第二，选择进行插值的数学关系。

第三，搜索落在窗口内的样点。以方框或圆框窗口为例。假定方框边长为 $2a$（若取图中所示的 2×2 网格，a 即网格或像元的边长），待插点坐标为（x_0，y_0），则位置坐标为（x，y）的样点若符合下述条件

$$x_0 - a \leqslant x \leqslant x_0 + a \quad 和 \quad y_0 - a \leqslant y \leqslant y_0 + a$$

将被“套”在方框窗口内。若窗口为圆框，设圆框半径为 r（参见图 4－9），注意

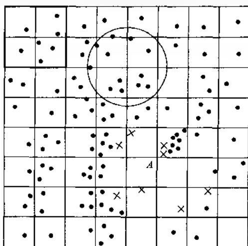


图 4-9 活动窗口插值示意

圆框实际上是一个与待插点等距离的范围，则满足下列条件

$$(x-x_0)^2+(y-y_0)^2 \leqslant r^2$$

的样点将落入圆框窗口内。

第四，基于每个待插点窗口所套住的样点数据，根据所选择的数学关系进行插值计算，以确定窗口中心，即待插的网格点处的空间变量之值。

注意，待插点为规则格网上的点，依次处理待插点，相当于在 x 或 y 方向移动一个网格步长。因此，很容易通过循环语句一个接一个快速地搜索每个待插点的窗口所套住的样点，并一个接一个地计算出待插点的值。在 GIS 实践中，不少场合对精度要求不是很高，采用活动窗口法形成规整的栅格数据，特别是数字高程的栅格数据，是一种非常简单易行的、有效的实用方法。活动窗口法也存在一些问题，如有时会有较大的数据质量损失，区域边界网点的插值需要特别处理等。还有样点搜寻也会遇到一些问题，下面再谈。

活动窗口法中借以进行插值的数学关系可以是多种多样的。根据所采用的插值数学关系的不同，活动窗口法可进一步分为若干种类型。下面就来介绍两种常用方法。

2. 局部函数法

局部函数法用一个数学关系来逼近或拟合待插点窗口所套住的样点数据，类似于上面的趋势面分析。不同的是，趋势分析法通常用于全局分析，因而特别讲究采用回归面；而这里是小范围插值，以逼近或拟合为主要目的，不强调趋势面、

局部函数法也习惯采用多项式函数。由于是小范围处理，样点数据有限，一般仅采用一次或二次多项式。当采取如上面式(4-5)所给出的二次多项式时，要求被套住的样点至少为6个，大于6个时采用最小二乘法。利用坐标为(x_0, y_0)的待插点窗口所套住的已知(x, y, z)的样点数据，基于式(4-6)所示的线性方程组，求解式(4-5)的6个系数。然后，将待插点的自变量值(x_0, y_0)代入式(4-5)，即可决定待插点(x_0, y_0)处的空间变量z值。

现在很多软件都有解线性方程组的现成功能，采用多项式拟合已较为方便。

3. 移动平均法

移动平均法也称为滑动平均法，它直接采用待插点窗口所套住的样点数据的简单平均值，或加权平均值，作为待插点(x_0, y_0)处的z值：

$$z(x_0,y_0)=\Sigma(a_i z_i),\qquad \Sigma a_i=1\quad i=1,2,\cdots,n \qquad (4-7)$$

式中，z_i和a_i分别为待插点窗口所套住之第i个样点的空间变量值和权重系数；n为待插点窗口所套住的样点数，当采取简单平均时，$a_i=1/n$。可见，移动平均法比上述局部函数法更为简便，而且不要求至少被套住多少样点(当然，太少也不好)。

在加权平均的情况下，权重系数a_i多采随距离减小而增加的函数关系。这是为了进一步体现理论假设，即空间位置上愈靠近的点，愈可能具有相似的特征，空间变量的值愈可能相近。事实上，局部插值方法用邻近的样点数据来插值，本身已经体现了距离近处权重大的原则；现在更进一步，在窗口所套住的邻近样点中，给予更近的样点以更多的权重。

采用随距离减小而增加的权重系数的加权平均常见形式是

$$z(x_0,y_0)=\frac{\Sigma(z_i d_i^{-r})}{\Sigma d_i^{-r}}\qquad i=1,2,\cdots,n$$

式中，d_i为待插点窗口所套住之第i个样点至待插点的距离；其余字母意义同上。r常取为1，即用距离倒数作为权重，或者，$r=2$，采用距离平方的倒数作为加权平均的权重系数。

移动平均法的主要问题是容易平滑掉区域变量的尖锐变化，其作用类似于遥感图像处理的低频滤波。例如，对不规则的高程数据进行移动平均，假定待插点恰是山尖，山尖的高程应当比周围任何样点都高；但是，采用移动平均法，即使用距离反比权重，所得结果也将低于周围最高的样点。同理，谷底处的移动平均插值也会有不合理结果。山尖和谷底处若采用式(4-5)所给出的2次多项式，结果就会合理得多。此外，在陡坡和断层等高程急剧变化的区域，移动平均法也可能使急剧变化模糊。

4. 搜寻策略

搜寻策略讨论活动窗口法中怎样决定搜索范围，以便有效地搜寻到更合理的计算插值的样点。这个问题是编写不规则数据向栅格数据转换的程序时会遇到的实际问题，关系到活动窗口法的插值结果，因而有必要进一步讨论。

首先，一种搜寻策略是固定搜索范围，例如，采用固定大小的方框或圆框窗口来进行搜索。固定搜索范围的优点是计算简捷，缺点是所套住的样点数可能很不均匀，有的窗口窗口套住的样点数很多，有的很少，甚至达不到必要的数目。为了达到必要的数目，必须先编写探察样点数的程序；但当窗口能保证最小样点数时，另一些窗口套住的样点数可能过多。

另一种搜寻策略时固定搜索数目。这种策略采用大小可变的窗口，以确保搜寻到固定的样点数目，例如，采用大小可变的圆窗，即距离待插点最近的 n 个样点数据来插值。固定搜索数目的策略也不难用程序实现，但运行比上一种策略费时，因为对每个待插点，都要进行全部样点的距离比较。更主要问题是这种搜寻策略会遇到样点串(cluster)的困难。例如在图 4－9 中，A 点附近有一样点串，这时采用固定搜索数目的办法，计算 A 点插值的样点将集中在一个狭窄的方向，这显然不合理。例如，在高程数据情况下，若样点串位于一个小高地，插值的结果肯定不能反映本应较低的待插点的值。

解决样点串困难的一个方法是按象限或半象限搜索。这种策略要求在进行局部内插时，每个象限或半象限都贡献同样多的样点。仍以图 4－6 中的 A 点为例。若采用象限搜索，可以规定在待插点周围每个象限分别取最近 n 个样点作为计算插值的样点。例如，取每个象限中的最近 2 点，共取 8 个样点，如图中 8 个“×”点所示，显然，取这 8 个样点计算 A 点插值，比取 A 点西北 8 个最近的密集样点，要合理得多。半象限搜索是把每个象限再按 45 度分成两半，每半个象限取最近几个样点。象限或半象限搜索相当于固定搜索数目策略的发展，虽然增加了一点编程的内容，但运行时间不会明显增加，因为两种策略都需要对每个待插点进行一次全部样点的距离比较，只是象限或半象限搜索的比较是分块进行罢了。

象限或半象限搜索是活动窗口插值的较好的、可实用的方法，但区域边界处的插值可能需要多一点特殊处理。

4.3.4 其他局部插值法

1. 克里金插值

克里金插值又称为统计优化法，一般而言，它不是用来直接进行插值，而是一种应用地理统计理论优化插值，特别是优化局部插值的方法，它可以为局部插值提供确定权重系数的最优方法，并能描述误差信息。克里金插值是唯一的

统计优化算法，由南非的克里金（D. G. Krige）在金矿勘探应用中提出，因得名，并得到 Matheron 改进。他们提供了确定权重系数的最优方法，从而使内插函数处于最佳状态，对给定点上的变量值提供最好的无偏估计，因而被广泛地应用于地下水模拟、土壤制图等领域，并成为 GIS 或某些地学专业软件地理统计插值的组成部分。

克里金插值基于区域性变量的数学理论。区域性变量理论认为，任何在空间连续性变化的属性是非常不规则的，不能用简单的平滑数学函数完全模拟，但可以用随机表面给予较恰当的描述。这种连续性变化的空间属性称为"区域性变量"。这样，该理论将空间变量分解为三部分：趋势或结构性成分，与空间变化相关的随机部分（区域性变量），以及随机噪声。就象沿小路爬山，总起来说是沿坡的趋势线上行，但中间也穿插着局部的上下波动（随机相关的高程），以及不时遇见的需要跨上的大石头（高程噪声）。区域性变量理论假定差异的稳定性和可变性，一旦结构性成分确定后，剩余的差异变化属于同质变化，不同位置之间的差异仅是距离的函数。

克里金插值包含几个步骤。首先，用一个数学函数来估算（整个区域的）趋势或趋势面。如果没有趋势函数，则采用平均值作为估算基础。区域性变量由趋势面和半方差给出。计算半方差的窗口范围或样点搜索区，如上述，随不同搜寻策略而不同。然后，用一个数学模型来拟合半方差图样点的变化，通常取指数函数或球面模型；但当半方差变化较简单时，可取线性模型。利用这些模型，就可以估算用于活动窗口加权平均，即式（4－7）的权重系数。采用这样权重系数的插值将是一种最佳的线性无偏估计，估计方差最小。

关于克里金插值的较详细讨论，可参见邬伦等《地理信息系统——原理、方法和应用》、龚健雅《地理信息系统基础》和张超等的《地理信息系统》。

由于克里金插值给出确定权重系数的最优方法，并能描述误差信息，它优于上面讨论的其他内插方法，具有普遍意义。克里金插值特别适宜于具有很好定义的趋势面的数据。但是，当用邻近样点来估计半方差较困难时，所用的模型就不完全合适，克里金插值就不一定好于其他方法。克里金方法在统计学上的优越性，也带来计算上一定程度的繁难性，特别当有大量点数据和格网点时。Burrough 曾评述了几种克里金法与其他方法的比较研究。一种克里金法与距离加权法（距离平方反比）的比较结果表明，克里金法的估算值只在一小块缺乏采样点的、有样点串的边缘区域，明显优于距离加权法。

总起来说，克里金插值应用于精度要求较高，同时样点数据有较好的趋势面的场合。在不要求这两个条件时，采取其他简便易行的方法，也未尝不是一种好的选择。例如，现在不少 GIS 工作进行三维模型处理时，主要以动态的、生动的形象显示为目的，不十分重视地形细节的小偏差，这时，为形成高程栅格数

据，不必采用克里金插值方法

2. 加密插值和配准插值

加密插值和配准插值都是在已有(规则或不规则的)网格数据的条件下，求取网格中间的点的插值。其中，加密插值可在三角网或方格网中进行；配准插值仅在方格网中进行，它是由于进行图像坐标校正时，校正后的图像与原图像的像元"错位"，以致校正后图像的像元属性不能直接取用原图像的像元属性值(新图像的第 i 行第 j 列的像元不能采用原图像 i 行 j 列的属性值)，而需要基于原图像中相应位置周围的若干像元的属性值来内插。这个工作称为"重采样"。有关重采样的详细内容将在遥感图像处理课程中讲解，这里，我们只需要记住一点就行了，即：这种坐标校正导致的配准插值的方法，与下述的方格网加密方法是相同的。

无论在三角网还是方格网中间插值，通常就以待插点所在网格(必要时加上周围网格)为加密单元，并采用低次多项式拟合。方格网加密多采用双线性内插法，即在网格的 x 和 y 两个方向都进行线性内插。参照图 4-10b，设格网的行、列方向即坐标轴方向，每个网格的边长分别为 Δx 和 Δy，待插点 (x, y) 所在网格的四角点上各有样点数据 z_{11}、z_{12}、z_{21} 和 z_{22}。为了内插出 (x, y) 处的 z 值，必须先在 x(或 y)方向上内插出与两个 y(或 x)值对应的 z 值，然后再在 y(或 x)方向内插出最后结果。因此，双线性内插实际上要进行 3 次线性内插。下面写出先在 x 方向上内插的双线性内插公式。首先，在 y_1 水平线(图 4-10b 的下横线)的 x 方向，我们有

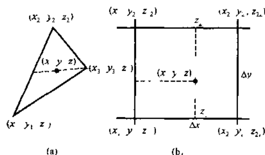


图 4-10 加密插值

$$(z_1 - z_{11})/(z_{21} - z_{11}) = (x - x_1)/\Delta x,$$

从而

$$z_1 = z_{11} + (z_{21} - z_{11})(x - x_1)/\Delta x$$

同理，在 y_2 水平线上，有

$$z_c = z_2 + (z_4 - z_2)(x - x_2)/\Delta x$$

然后,基于求得的 z_r 和 z_c,沿着图4-10 b 中的纵向虚线,在 y 方向上内插,可得双线性内插的最后结果,即待插点(x, y)处的 z 值:

$$z = z_r + (z_c - z_r)(y - y_r)/\Delta y$$

从上可见,双线性内插是基于一个方格的四顶点之值来进行内插。在遥感图像处理的配准插值中,有时还采用包围待插点的 9 个方格共 16 个网格点之值来进行插值。

在三角网加密的条件下(参照4-10 a),可采用简单的线性关系

$$z = ax + by + c$$

通过待插点所在三角形的三顶点(x_1, y_1, z_1)、(x_2, y_2, z_2)和(x_3, y_3, z_3),来求解上式中三个系数 a、b 和 c,从而内插出(x, y)处的 z 值。三角网加密也可以采用双线性内插,但相对麻烦一些。这时可以过三角形某顶点和待插点作一辅助线,交于对边,如图4-10 a 所示。先沿对边,内插出交点处的 z 值;然后基于这个 z 值和该顶点的 z 值(图中为 z_3),内插出待插点(x, y)处的 z 值。

3. 样条函数

样条函数主要用于区域内小片区块数据空缺或局部数据修改,为此,它一般采用多项式模拟局部样点的分布,并使这种局部模拟在该局部边界与原有的数值自然地衔接和拟合,从而可能较真实地反映空间变量原有的变化趋势。

为什么称为样条函数呢? 原来,手工绘图时,借助于一种称为"曲线规"的工具,可一段一段地绘出拟合样点的平滑曲线,这种分段曲线就称为样条。曲线规绘出的曲线近似于分段的三次多项式函数,并在交接处与原有曲线比较自然地衔接和拟合。参见图4-11,图中两条实线表示原有的曲线,中间有一段数据空缺,需要拟合一段样条,如图中粗虚线所示。该样条的两头都与原曲线较好衔接,实际绘图时,在两边拟合好后绘出中间空缺的一段。由此可见,补充或修改的样条具有一定合理性,如果原空缺处为山顶部分,样条可能近似反映原有的山顶曲线的形状,这种效果是活动窗口等方法所达不到的。



图4-11 样条示意

样条函数模拟就是基于上述原理,样条函数在数学上相当于灵活曲线规的一个分段函数。当然,图4-11 所绘样条为一元曲线,而空间变量的样条插值是二维曲面内插,采用二元样条函数(通常取2元3次多项式)。为保证在边界处与原有数值自然衔接,样条函数在连接处有连续的、与原有样点数据相等的一阶和二阶连续导数。样条函数只进行局部区块的拟合,用以补充或修改局部区块的空间变量分布曲面,而不用理会或处理不涉及局部区块修改的其他部分

的数据，这是趋势面分析方法做不到的。由于是分段函数，样条函数还有插值速度快的优点。样条函数的主要缺点是，样条内插的模拟虽然看上去比较满意与合理，但其误差不能直接估算

样条函数在 GIS 中，除了可用于二维的空间数据插值外，一元的曲线样条也常用来修改矢量地图中的曲线形状。例如，一端曲线中间点过少，折线形态明显，就可以用一段曲线样条增补中间点，使该段曲线看上去不仅平滑，且与两边自然衔接。

§4.4 地理空间数据的三维处理

4.4.1 空间数据三维处理的基本概念

1. 空间数据三维处理的两种基本思路

地理空间数据三维处理，特别是三维地形表达和分析，在传统的地学和 GIS 领域一直受到高度的重视。作为很多地学分析、表达和应用的基础数据和手段，它不仅具有重要的基础研究意义，而且具有很大的应用意义，广泛应用于农、林、牧、水利、交通、军事、区域规划和建设等领域，例如公路、铁路、输电线的选线、地面工程的选址、军事制高点形选择、土壤侵蚀和土地类型的分析等。近年来，随着数字地球时代的到来、三维 GIS 技术成熟化和科学可视化要求的不断提高，三维立体表达和分析在 GIS 中受到愈来愈多的关注，成为 GIS 领域最重要的研究领域之一。

空间数据三维处理的基本思路有两种。一种称为真三维处理。这种处理将三维空间的地理位置(x ,y, z)看作自变量，而将空间变量作为第四维的因变量。为此，这种方法要求建立三维空间数据结构，并基于三维数据结构进行立体表达和分析(回顾见 §2.4.3.4)。真三维处理的研究具有重要意义和前景，而且目前已有不少的努力，如八叉树数据结构和三维矢量数据结构等，但由于难度较大，目前研究进展不大。

三维处理的另一种基本思路，称为2.5 维处理方法。目前 GIS 的空间数据三维处理，包括本节将要讨论的三维处理，皆指这种方式，对此将不再说明。2.5维处理类似传统的美术技法，它基于二维数据结构来表达三维形象。在这种处理中，空间变量 z 被看作是二维的地理位置(x, y)的二元函数

$$z = f(x, y)$$

即一个随自变量(x, y)而起伏变化的二次曲面。Z 可以是任何空间变量，如高程、地价、土地利用、土壤类型、地貌特征、岩层深度和污染指数等；二次曲面 $z=$

$f(x,y)$就是该变量Z的区域分布。这种分布的立体形象也可以基于透视原理，像美术画那样在平面上表达。图4-12是一幅形象地表现了某市2000年基准地价区域分布之起伏形态的三维图。

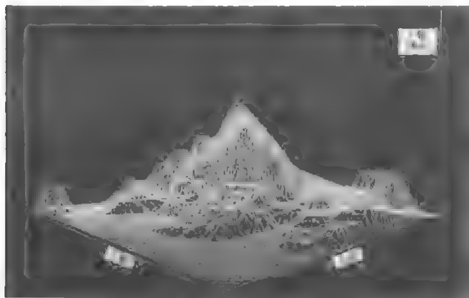


图4-12 某市基准地价(2000)

变量Z虽然可以是任何空间变量，但有一个变量对三维处理来说是最基本、最重要的，那就是高程。这是因为，如果空间变量Z是高程，那么，$z=f(x,y)$就是地球表面真实的起伏地形；而地形的三维（立体）表达和分析，是空间数据三维处理的本源。事实上，其他空间变量的三维表达，例如图4-12，其实是借助于GIS地形表达的手段实现的。实现这一点不难，只要在GIS软件的三维地形模块中将高程变量的输入，代之以输入其他空间变量就行了。由此缘故，下面我们将仅讲解地形的三维（立体）表达和分析。

很久以来，人类已采取了各种可能的手段和方法进行地形立体表达与分析，其中大部分现已为GIS所借鉴。因此，我们首先来梳理一下人类进行地形立体表达的各种方式，这对学习GIS三维空间处理有帮助。

2. 三维地形表达的传统方式

三维地形表达的传统方式有多种，其中一种是数学模拟方法。这一方法用数学方程或某种数学函数的级数来模拟高程变量在整个区域的分布，也就是说，给出有解析表达式的地形曲面。数学模拟常采用多项式函数，也可以

采用傅立叶级数等。有了描述高程变量空间分布的解析表达式,不仅能把握整体地形,而且可以从中求取地形特征点、线的信息,如山巅、等高线、山脊和谷底线等。但是,数学模拟方法在实践中往往不现实。这是因为,地表形态极其复杂,地形模拟涉及的数据量通常又太大;而且,由于高程数据的特点、比例尺和模拟的精细度等很多条件的限制,要得到满意的模拟也很不容易。

常见而有效地表达地形立体起伏的方法是图形图像法。其中,最常用的是采用等高线法。该方法把地面上高程相等的各相邻点所连接而成的闭合曲线,垂直投影到平面上,形成等高线图。

在等高线图的基础上,可进一步采用分层设色法,即根据地图的用途、比例尺和区域特征,将等高线划分为一些层级;每一层级的区间内普染不同颜色(也可用不同灰度),以色相、色调的差异来表现地势高低。这种分层设色的等高线图能更加明显地表达地势状况,使地图有了一定的立体感。

再进一步,可采用晕渲法。晕渲法又称为阴影法,它用反映一定光照下之各地明暗不同的色调,来表现地形的起伏形态。晕渲法类似于美术中的油画,立体表达的效果非常好。传统的晕渲法通常只用于垂直投影的地图上。

传统技法中有一种GIS未采用的立体地形表现手段,这就是写景法。写景法又称为鸟瞰法或透视法,它是基于透视原理,采用类似美术中的线条技法(回顾见§2.1.4.4和5)。写景技法很早就为古人采用,例如用眉毛似的线条表现远山等。近代写景法采用直观的透视符号,表示基本地貌类型及其各种地貌形态。但应注意,现代GIS虽然没有直接采用传统的写景法,但却更有效、更科学地应用了透视原理,发挥其功能。在矢量数据或线条表达方面,GIS用透视原理来立体表现方格网和三角网的网线起伏,图4-12就是方格网起伏的立体透视图。在栅格表达方面,GIS集晕渲法与透视法之大成,不仅透视计算每个网格的立体透视位置,还计算每个像元在一定光照下的明暗、色调,形成高度仿真的立体透视晕渲图,极大地扩展了传统地图的地形起伏的立体表现方法(详见本节后文)。

三维表达地形的其他方式还有特征线方法,也称为线模式方法。线模式方法包括等高线和其他特征线的表达,如水平线、垂直线和典型地形线等。典型地形线指山脊、谷底、海岸和坡度变换线等。除了线模式外,点模式法也是表现地形的一种手段,典型的点模式表达地形的例子,是地形图上众多的高程点。

3. GIS三维处理的主要内容,DEM和DTM概念

GIS三维处理是传统地形表达的延伸,现在已包含很丰富的内容。本小节先给出GIS三维处理总体轮廓的描述,作为后文从各种角度进一步展开讨论的纲要。从大的方面看,地理空间数据的三维处理主要包括四方面的工作(参见图4-13)。

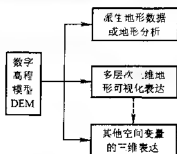


图4-13 GIS三维处理主要内容

• 建立数字高程模型DEM（digital elevation model）。数字高程模型DEM是以(x,y)为自变量的高程z数据的有序集合。由于地表极其复杂，用解析表达式来模拟不大现实，GIS在三维建模时，通常采用离散的高程数据，这些数据的样点按一定规则布设在整个研究区域。常用的DEM有两种形式：一种称为高程矩阵，高程数据布满覆盖整个区域的方格网的网格，相当于高程的栅格数据；另一种DEM的高程数据布满覆盖整个区域的三角网网点，实际上就是TIN数据(见§2.4.3.2)。

数字高程模型把研究区域分成众多的、有高程数据的地面单元。在高程矩阵的情况下，最小地面单元是方格网的网格；在TIN的情况下，最小地面单元是三角形。基于这些地表单元，就可以进行三种进一步的三维处理工作(图4-13)。可见，数字高程模型是GIS三维处理的基础。关于数字高程模型及其生成的进一步内容，将在本节中的§4.4.2讲解。

• 计算派生地形数据或地形分析。基于DEM可以计算每个地表单元的坡度、坡向、粗糙度和可视域等地形因子，以及集水线、山脊线和谷底线等地形轮廓线。这些数据称为由DEM派生的地形数据。从这些派生数据的名称就可见，DEM派生数据的计算实际上就是进行地形分析；而地形分析是一种GIS空间分析。因此，DEM派生数据的内容有时与三维处理一并讲授，有时在某些GIS书中被放在“空间分析”一章。本书采用前一种方式，本节中的§4.4.4将讲解这部分内容。

• 各种层次的三维可视化地形表达。基于DEM还可以计算一定观察角度，一定光源条件下每个地表单元的明暗、色调和透视关系等，这就是DEM基础上所做的另一大类工作：在二维画面上可视化地表达三维(立体)地形。GIS三维处理中，可视化地形表达有多种层次，如图4-12那样的透视立体图、照片般更进一步的彩色立体图和更高级的三维表达等。本节中的§4.4.3及下一小节，将要进一步介绍有关的三维可视化地形表达技术。

• 其他空间变量的三维可视化表达。DEM及三维可视化地形表达技术，如前所述，可以借用来可视化表达其他空间变量的空间分布形态，就像图4-12那样。这是DEM之上所进行的第三类GIS三维处理工作。

基于上述分析，现在可以较好地解释DTM(digital terrain model)概念了。DTM是通常并列于DEM的另一个最常见基本概念，任何GIS书籍讲三维处理时都会提到它。但DTM在各处的译名和用法不像DEM那样一致。在翻译上，

有些书上将 DTM 译为数字地形模型(直译);另一些译为数字地面模型。在对 DTM 理解上,有的作者偏重地形含义,即偏重上述第二类工作,例如,将 DTM 定义为"空间地形数据集合的统称,是带有空间位置特征和地形属性特征的数字描述";另一些作者认为 DTM 是"描述地面特性的空间分布的有序数值阵列",这种解释可以涵盖任何空间变量的三维表达,即包括上述第三、四两类工作(甚至可以适用于 DTM 以外的场合)。有的书上先定义 DTM,然后再将 DEM 定义为"空间变量或地面特性为高程的 DTM"。

我们认为,DEM 与 DTM 名词产生于 GIS 发展的早期,那时地形表达和分析的内容还不是很多,DTM 原来的用意是涵盖除 DEM 基础模型外的三维地形处理。现在,GIS 三维处理的内容已大大扩展了,DTM 英文原定义已不能覆盖 DEM 外的 GIS 三维处理的内容,大家对 DTM 的理解出现见仁见智的局面,也就不奇怪了。

4. 现代三维地形表达的 5 个层次

GIS 领域对三维处理技术进行分类的工作,总起来说不多。柯正谊等曾按空间结构形式将 DTM 分为 7 类:规则格点(格网)数字地面模型、散点数字地面模型、等值线数字地面模型、曲面数字地面模型、线路数字地面模型、平面多边形数字地面模型和空间多边形数字地面模型。作者认为,现在有必要再从一个角度,即技术复杂性的角度,区分不同层次的三维表达技术。这是因为近年来,GIS 三维表达技术发展得非常迅速,一些过去不容易实现的"高难度"三维表达,现在已开始普及;同时,更新、更高级的 GIS 三维表达技术又不断涌现。在这种形势下,作者在自己的科研和教学实践中,提出了 5 类不同技术层次的三维表达技术。下面给出这 5 种类型,读者可从中顺便了解近年来 GIS 三维表达技术的发展。

- 第一层是数字高程模型。定义为以(x,y)为自变量的高程 z 数据的有序集合,包括方格网数字高程模型(高程矩阵)和 TIN 模型等。DEM 是实现较高层次三维表达的基础。对这一层本身进行可视化地形表达,可采用按高程(或等高线)的分层设色法(参见图 4-14 和 §4.4.1.2)。
- 第二层是数字地形模型。这一层的特点是已具有照片般的可视化地形,它已应用透视原理,通过明暗或色调处理,可鲜明表达出任意鸟瞰角度和光源条件下的地形起伏,场景可随观察角度而移动。但这种地形模型的地面上尚空无一物,看上去就像月球和火星表面那样空旷。因此,这一层采用英文 DTM 非常恰当。但若考虑到一些 GIS 学者对 DTM 一词的用法已有传统的理解,也可以考虑将此层命名为数字起伏模型(digital undulation model 或 digital relief model)。
- 第三层是数字地面模型,这一层数字地形模型的地面上,叠加上真实的

或仿真的地表或其地物，铺设仿真地面的材料可以是工射航片、卫星影像和线划地图等，还可以按照地图实际情况添加仿真地物，如按图造房子或种树等，这些地物可通过额外的建模制成，并被植入到地表上。因此，这一层称为数字地面模型非常恰当，英文可采用 digital ground model，简称 DGM。若考虑到中文"数字地面模型"一语用法的传统习惯，这一层也可称为电子沙盘。电子沙盘是一个现成的用语，已广泛用于军事、水利和区域规划建设等领域。电子沙盘相当于传统的沙盘，是计算机内能从任意鸟瞰角度和光源条件下观察的仿真地面模型（当然，仿真的内容因专业不同而异），正好相当于我们这里说的数字地面模型。

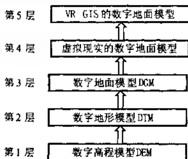


图4 14 现代地形表达的5层次

● 第四层是虚拟现实的数字地面模型。虚拟现实（virtual reality，缩写为 VR）是一种近年来十分活跃的计算机技术领域，是多种计算机分支技术与人类行为学等在更高层次上的集成和渗透，是多媒体技术发展的更高境界；它为用户生成逼真的、具有视觉、听觉、触觉等效果的可交互的动态世界。除交互性外，虚拟现实技术最主要的技术特征是沉浸感和身临其境感，并使用户在这种"全身投入"的感觉状态中，激发新的想象（详见 §7.2.1.1）。虚拟现实的数字地面模型，就是要求至少部分具有虚拟现实技术特征的电子沙盘，其英文可写为 digital ground model with virtual reality，简写为 VR－DGM。

在当前条件下，完全达到高档次虚拟现实技术的要求还不现实，即使在发达国家，虚拟现实技术要求的设备，如头盔和数据手套等也并未普及。为此我们认为，虚拟现实的数字地面模型至少要达到两个条件：第一，地表环境中应有虚拟的地物，例如规划中的、想象的地物；第二，必须能够按任意指定路线让用户动态漫游，并辅以多媒体技术，使用户至少能有身临其境的感觉。

● 第五层为虚拟现实/GIS（或 VR－GIS）的数字地面模型。上述第四层的技术档次虽然已不低，但它还未达到 GIS 的基本要求，即还不能进行地理空间查询。而为了达到 GIS 这一基本要求，上述虚拟现实的数字地面模型必须实现与数据库的挂接，至少应使用户在身临其境的三维动态漫游中，随时能查询到它所到达的河流、村庄等地物的地理信息。这才是虚拟现实-GIS 的三维地面模型，也是当前最高的层次的 GIS 三维表达技术。作者领导的浙江省 GIS 重点实验室业已自主开发了这一层次的三维表达技术。虚拟现实－GIS 层次的数字地面模型的英文表达，可简写为 GIS－VR－DGM。

4.4.2 数字高程模型及其生成

鉴于数字高程模型 DEM 在 GIS 三维处理中重要的基础性意义，有必要较详细讨论之。下面先讲解数字高程模型生成的一般思路，然后再分别具体讨论两种主要的 DEM，即高程矩阵和 TIN 的生成及其特点。

1. 数字高程模型生成的一般思路

数字高程模型及其生成的一般思路如图 4-15 所示。首先，在图中 DEM 的右端给出数字高程模型的两种主要类型：高程矩阵和 TIN。图中可见，不论哪种 DEM，都一般以不规则分布的高程样点数据为直接加工材料，但加工的途径不同：从不规则分布的高程数据获取高程矩阵的主要途径是内插；获取 TIN 则可以直接通过组合。

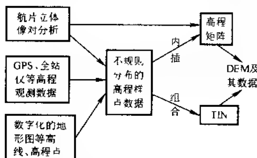


图 4-15 数字高程模型的生成

图 4-15 的左部三个图框给出不规则高程数据的几种主要来源。一是航片立体像对分析。航空摄影测量已有非常成熟的技术，包括手工的、半自动和全自动的方法，通过航测相片之像对大批量地求取地面点的高程数据（等间隔采样可直接得高程矩阵），现代全数字摄影测量技术进步很快，大有取代其他测量方法的势头。高程数据的第二种主要来源是利用测量仪器在野外采集。野外采集高程的技术也进步得很快，早些年所用的水准仪、经纬仪和测距仪等设备，现在已普遍为 GPS 加全站仪所取代。

尽管测绘技术有了如此进步，但对于无法从测绘部门获取高程电子数据的大多数 GIS 用户来说，第三种获取途径更重要，也更常用。这就是通过地形图等高线和高程点的数字化来获得不规则高程数据。§2.4.3.1 和 2 说过，用扫描矢量化或数字化仪手扶跟踪一条条等高线，并且在输入 GIS 中后进一步赋予其高程属性，实际上就是采集到表达各等高线的、具有一定高程 z 值的一串串 x, y 坐标点。这正是不规则分布的 (x, y, z) 数据之集合。

有了不规则分布的高程样点数据，就可以分别通过内插和直接组合的途径，建立高程矩阵是和TIN。下面具体讲解这两种最主要的数字高程模型的特点及生成

2. 高程矩阵的特点及其生成

高程矩阵又称为格网DEM或Grid，是以高程为属性值的栅格数据 高程矩阵以规整格网的单元，或像元，作为表达高程起伏的基本地表单元或地形单元。它具有栅格数据的优点，如结构简单、规则，采取矩阵数组形式方便计算机处理，以及有利于计算等高线、坡度、坡向、山地阴影、描绘流域轮廓等。这些优点使高程矩阵成为DEM最通用的形式。当然，高程矩阵也会有栅格数据的数据量大和精度受分辨率限制等一般缺点；特别在地形表达上，均匀网格的高程矩阵的内部分辨率(像元大小)一成不变，以致在地形少变、缓变的地区，数据可能大量冗余，而在地形剧变、跃变的地区，网格又可能太粗，不能反映地形细节，难以精确表达复杂的关键地形，如山峰、洼坑、隘口、山脊和山谷线等，影响地貌学分析。

生成高程矩阵的实用、常见的方法是在选定合适分辨率的方形网格的基础上，采用活动窗口插值法，将不规则分布样点的高程数据插值到规整网格点上，即形成高程矩阵。插值时常用距离反比加权平均法，必要时采用克里金法；网格数据加密时常采用双线性插值。当然，各种插值方法的选择及具体实现过程中的参数选择或调整，应具体情况具体分析，视地形和数据的实际情况而定 有研究表明，等高线数字化来源的高程数据经内插产生的DEM，其品质随内插法、等高线的质量和地形特性等因素的变化而异。没有一种对任何情况都完美的方法，§4.3.3曾评述过每一种插值方法的优缺点，可参考之。

除上述插值方法外，生成高程矩阵还有其他方法，最先进方法是全数字摄影测量法。现代全数字摄影测量工作站已能从航测相片之像对直接快速地生成高程矩阵，并能随地形的复杂化自动增加网格数量，提高分辨率，分块绘制等高线图。

生成高程矩阵的最原始方法是人工网格法，该方法将合适的格网蒙在地形图或不规则高程样点上，逐格读取或插值计算高程值，再输人计算机中形成数字高程模型。这种方法虽然原始，但对GIS初学者和某些普通用户来说仍有一定的学习意义。

3. TIN模型及其生成

TIN(triangulated irregular network)，即不规则三角网，是直接基于不规则高程样点进行组合，将离散高程样点每三个邻近点联结成三角形(当然，联结的方法还很有讲究，详见下文)，以连续的三角面形成的数字高程模型。不规则三角网以三角面作为表达高程起伏的基本地表单元或地形单元。

TIN 的数据结构已在§2.4 3.1 和 2 中讲过。我们已看到，TIN 数据兼具栅格和矢量两种数据结构的特色。一方面，它像栅格结构那样，用一个格网覆盖整个地区，试图用格网的每个单元及其属性（坡度、坡向和明暗色调等）组合起来表达地表形态；另一方面，TIN 的地面单元不是方形，而是不规则的三角形，不能以矩阵形式运作，因而不得不借助于矢量方法来描述三角形单元的点、线及其拓扑关系。当然，TIN 所涉及的拓扑特性不多。图 2－18 曾给出 TIN 数据的一种描述。在那里，只需两种拓扑关系，即每个三角形由哪三点组成，以及与哪两个三角形邻接；再加上每个结点，即数据样点的(x, y, z)数据，就能够把握三角网及其每个点和三角形的相互关系。

由于需要运作三角形单元的点、线及其拓扑关系，TIN 模型在计算上没有高程矩阵简便；但是，它却有两个重要的高程矩阵没有的优点：

第一，从不规则样点数据构筑高程矩阵，需要内插，而内插总是有误差的（§4.3 1.1）；TIN 却无需内插，因而避免了内插处理的精度损失。

第二，在 TIN 数字高程模型中，三角形地形单元的形状和大小取决于不规则样点的分布：样点密集的地方三角形小而密，样点稀疏的地方三角形大而疏（参见图 2－17）。而大多数实际测量和等高线数字化获得的高程数据，都有样点分布疏密随地形变化而异的特点：在地形变化复杂、频繁的地方，样点一般较都多。因此，从不规则样点数据构筑 TIN 时，TIN 的三角形密度和大小常常能自然地反映地形的复杂变化，在变动频繁复杂的地方，三角形较小而密集；在地形平缓、样点较少的地区，三角形疏而宽。TIN 模型这种内部网格密度或分辨率可随地形特征而变化的特点，使它能一定程度地避免高程矩阵的缺点，一方面在地形缓变的地区减少数据冗余，另方面又可能反映地形的较细致变化，表达较复杂的关键地形。

从不规则样点构筑 TIN 有多种方法。一种最容易想到的方法是搜寻最近点。具体做法是：首先取其中任一点，并寻找与此点距离最近的点，连接这两个点构成第一边；然后在其余各点中寻找与这条边最近的点，找到后即构成第一个三角形（其数据进入图 2－18 所示数据表）。再以这个三角形新生成的两边为底边，分别寻找距它们最近的点构成第二个、第三个三角形。依此类推，直到把所有的点全部连入三角网中。

有两个最基本的搜寻原则需要注意：

第一，在三角网中共享一条边的三角形最多两个。因此，在以一边为底边进行扩展时，应先判断此边是否已被两个三角形共享。

第二，以初始的第一边为底边搜索第 3 个顶点时，应该在第一边的左右两边都搜索，形成左右两个三角形（除非第一边在区域边界上）。接下来以新三角形的两个新边为底边扩展时，则只需在一边扩展；但整个扩展过程要统一，或左

边或右边。注意,线之左和右取决于线的方向(参见§2.2.2.3);而在TIN数据中,线的方向由三角形顶点序号体现。例如在图2-18中,三角形 B 在nodes表中排为1,2,3,说明三角形 B 三条边的方向为顺时针旋转;若排为2,1,3,三条边的方向为逆时针旋转。由此可见,统一在左向(或右向)搜索,就是要求TIN数据三角形顶点排序规则的统一,或顺时针,或逆时针,二者必居其一(图2-18中的nodes表尚未排好)。

为了提高搜索效率,可采取分片的方法,避免每一次都搜索整个区域。例如,将区域划分成适当数量的正方形格网,并对每个三角形所在网格记录索引,当基于一条底边搜索第三点时,可首先确定底边之中点 P 所在网格,搜索此网格内所有点,必要时再搜索此网格周围的8个网格。

自动生成的TIN数据在山脊线、山谷线和断裂线等特殊地形线处,有时需要一点人工干预,因为某些TIN三角形可能跨越这些地形线,三角形平面不能反映这里的凸、凹地形或高程的跳跃变化。这样的三角形需要人工调整。

4. 关于TIN生成方法的进一步讨论

上面只讲了TIN生成的一般方法。事实上,不规则三角形构网方法有多种,不同的方法构网的结果可能不完全相同,因而还有一个优化选择的问题。如图4-16所示,图a中的离散点可能构成b、c和d不同的网形。不难理解,三角网的建立应基于最佳三角形的条件,即应当尽可能保证每个三角形是锐角三角形或三边的长度近似相等,避免出现图4-16 c所示的过大的钝角和过小锐角。因此,上述基于一条底边搜索最近第三点的方法中,"最近"的含义还应明确。例如,一种规定是,"距离最近"指第三点到底边的中点距离为最短。这种算法较简单,也能有效避免过大钝角和过小锐角。另一种常用的方法基于角度最大原则,要求搜寻对底边构成张角最大的第三点。具体方法这里就不介绍了,有兴趣者可参见陈述彭等《地理信息系统导论》和龚建雅的《地理信息系统基础》。

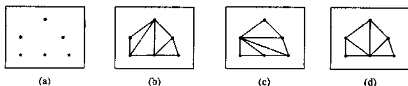


图4-16　不同的构网方式及其结果

还有一种非常有效的令人感兴趣的方法是狄洛尼(Delaunay)三角网。Delaunay三角网是泰森多边形或Voronoi图的对偶,我们已经在§2.4.3.2的小字内容中讲过。从图2-17可见,由泰森多边形的内参考点连成的三角网是比较理想的,能避免过大钝角和过小锐角。

为此，我们可以基于不规则高程样点，先生成泰森多边形或Voronoi图，然后将每一泰森多边形内的参考点与相邻的泰森多边形内的参考点相连，形成狄洛尼三角网。

从不规则样点构建泰森多边形，可以通过矢量算法，也可以用栅格方法。矢量算法直接基于泰森多边形的定义，即利用相邻样点连线的垂直平分线及其交点来构筑泰森多边形。

狄洛尼三角网也可以不经过泰森多边形或Voronoi图直接构建。这是因为，狄洛尼三角网有一个特性：每个三角形形成的外接圆都不包含有其他参考点。利用这一个性质可以直接构成狄洛尼三角网。

4.4.3 三维可视化表达的某些技术步骤

本节讨论在DEM基础上进行三维可视化地形表达的问题。上面曾讲过，现代地形表达技术有5种层次。本节仅介绍第一至第三层次的一部分内容。介绍中我们不作数学展开，而着重整体性的基本思路，希望能给出一个较完整的轮廓。第四、第五层次涉及虚拟现实的三维表达技术，非本课程基本要求，可参见本书第七章有关内容。

1. 数字高程模型的可视化

数字高程模型是基础层。由于DEM只是高程数据的有序集合，在可视化表达上，它只能采用分层设色的方法来表达地形分布。图4-17就是一幅用不同灰度表现不同高程的高程矩阵的DEM图像。从该图可见，这种方法虽然是一种可视化地形表达方法，但它缺乏立体感。



图4-17 浙江省1:10万数字高程模型(高程矩阵，局部)

如果在分层设色法中不是采用灰度表现，而是采用颜色，效果会稍好一些。但应注意，如此表现DEM时，要特别注意所设的颜色符合地图学一般习惯，如平原、丘陵和山地应分别采用不同层次或深浅的色彩过度适当的绿色、黄色和

褐色等。

2. 一定观察角度鸟瞰的透视立体图

要实现立体感较强的三维地形表达，必须基于DEM数据，应用透视投影原理，建立透视立体图。透视立体图模拟从某个视点，通过一个窗口（或一图框）观察世界的立体形象。图4－18就是一种表现方格网的透视立体图。透视立体图可以有很大的灵活性，可以根据不同的需要，任意变换视点和窗口使立体图形转动；或进行局部缩放，或改变高程的放大倍率以夸大立体形态等。透视立体图为研究地表形态提供了强有力的研究手段。

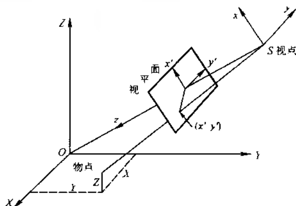


图4－18 透视立体图的制作原理示意（物点在原空间坐标系的坐标为X,Y,Z，在观察坐标系坐标为x,y,z，在视平面上的坐标为x',y'）

从数字高程模型到平面上的一幅透视立体图，其本质就是一个透视和投影变换。透视变换的方法很多，这里只讲解各种方法所共同的基本原理。如图4－18所示，设实际的空间坐标系或用户坐标系为(Z,Y,Z)，制作透视立体图的某数字高程模型数据集基于该坐标系。为了制作透视立体图，必须指定透视观察点和观察窗口，前者常称为视点；后者可作为今后输出画面的图框，或屏幕范围。视点和观察窗口分别如图中S点和平行四边形所示；观察窗口所在的平面称为视平面（也称透视平面或像面），与主视线或主光轴方向垂直。实际地点，通常就是DEM中的一个数据点，可称为物点。透视变换的关键就是求取物点未来在视平面上的投影位置。基于上述原理，透视立体图的制作一般分为三个主要步骤进行。

第一步是建立一个辅助坐标系来进行透视观察。该坐标系一般可称为观察坐标系，如图中(x, y, z)所示。其坐标原点建立在视点S上；一个坐标轴，通常取z或$-z$轴，为主视线或主光轴方向，一般指向原空间坐标系的原点。然

后，通过三维坐标系的线性转换公式（可查阅任何数学手册），计算出 DEM 中坐标为（X,Y,Z）的所有物点，在观察坐标系中的坐标（x,y,z）。为了今后更灵活的显示，原坐标的 Z 值可取为 $m(Z-Z_0)$，这里减去一个 Z_0 值是为了更好显示相对高程；m 为高程放大倍数，今后用户可指定不同的 m，增大或减小所显示的地形起伏的幅度。

第二步，在观察坐标系中设定观察窗口的参数，如窗口的尺寸和视平面距视点 S 的距离，这些参数的默认值可参照被观察的数字高程模型的范围来设定。设视平面上的平面坐标系为（x',y'），x'和 y'轴一般分别平行于 x 和 y 轴。由于现在所有物点在观察坐标系中的坐标（x,y,z）都已知，很容易通过简单的投影变换，计算出物点在视平面上的坐标（x',y'）。如果所有物点，例如高程矩阵的所有网格点的（x',y'）都计算出，那么，就可形成如图 4-12 所示的透视立体图。

第三，进行消隐处理，也称为可见面识别，就是在三维表达结果中消除地形地物中被遮蔽的部分。体现遮蔽关系也是透视原理和三维表达的重要内容。消隐处理的基本思想是，如果两个地物在视平面上的投影重叠，那么，比较两个地物离视点的距离（即 z 坐标），较远的地物的重叠部分应当被消隐，不出现在透视图上。消隐处理的原理虽然简单，但实际处理却比较复杂，主要是计算量太大。尽管已开发了不少可见面识别的算法，但没有一个对所有情况都是最好的。

3. 一定观察角度、一定光源条件下的灰度或彩色立体图

三维表达的另一种重要手段是借助于一定光线下图元的明暗或色调来表现。明暗的程度称为灰度。图 4-17 就是一幅黑白灰度图，只不过那里的黑白表现的是高程值本身；如果表现光照下的明暗和阴影，就会有很强的立体感，如同很多黑白的照片和电影一样。当然，用彩色色调，即有色彩的明暗调子更有表现力。用图元的明暗或色调来表现立体形象的图，分别称为灰度立体图或彩色立体图。为简单起见，我们仅以灰度立体图为例来讲解基本思路。然后，读者可基于彩色叠加原理（见 §2.3.4.3），推想彩色立体图的形成。

灰度立体图的制作通常包括下列步骤：

第一至第三步骤为上面已讲过的内容，它们是：第一，将地面分割为足够小的地面单元。通常采用三角形面片性能较优。第二，进行如上所述的透视投影处理，确定一定观察角度下的透视形象。第三，进行可见面识别，即消隐处理。只是现在不仅要消除从视点处观察的地形地物中被遮蔽部分；而且还要判断主要点光源（如太阳）的光线被另一些地物遮蔽而射不到的地物，即地形的阴影部分。这是下一步明暗处理所需要的。

第四，明暗处理，即依据光照模型计算可见地表单元的亮度。亮度是地表

单元的光辐射强度,除非本身发光,地表单元辐射是反射光。根据物理学规律,反射光可分为两部分:镜面反射和漫反射。理想的镜面反射的反射光只在反射角一个方向反射(入射角大于0);而漫反射可射向表面以上各个方向。无论哪种,反射光强度皆与地面单元的反射特性,如构成物质和粗糙度等有关。实际地面反射是两种反射的结合体。从视点处所观察到的地面单元的亮度是三种成分的反射光的合成:一是主要点光源(如太阳)的镜面反射光;二是主要点光源的漫反射光;三是环境光的漫反射。环境光包括天空大气散射的光和来自周围地表经多种反射形成的漫反射光。环境光及其漫反射一般可以看作是各向同性的,这将大大简化光照模型。上述三种成分的反射光中,第一、二两种在主光源阴影处将为零,主光源阴影是主光源因被遮挡而达不到的地方。但阴影部分的地表单元并不全黑,因为还有环境光的漫反射光。此外,明暗处理中有时还考虑距离因子的作用。

可见,明暗处理就是依据物理学定律,给定某些简化条件,建立光照模型,从而计算出从视点处观察的每个地表单元的亮度。

明暗处理是一个关键的步骤。明暗处理完毕后,一幅立体感很强的画面就会跃然纸上。前面的第三步骤形成了图 4-12 那样的透视立体图,但透视立体图的每个单元没有明暗和色调的差异。例如假定图 4-12 是一幅网格足够细腻的地形图,给定光源并经过明暗处理,每个网格单元显示出不同的亮度,该图的立体感将明显增强。

明暗处理完毕后,还有些细节需要修饰。下面五、六两步骤就是做这样的工作。

第五,地表面片的明暗过渡。当面片(地表单元)比较大时,如果每个单元只取一个亮度值,画面上可能表现出灰度的明显不连续,不能较好表达明暗均匀过度的连续表面。因此,有时还需要在面片内部赋予不同的亮度值。例如,一种方法是,确定三角形面片每个顶点处的灰度,再通过灰度内插而确定面片内部点的灰度。

第六,纹理处理。为了更细腻表现地表特征,必要时还要添加表面细节,即纹理。纹理处理分为两类:一类称为颜色纹理,主要用来表现表面较为光滑但有纹理图案的物体,如刨光的木材等。生成颜色纹理的一般方法是纹理映射,即预先定义纹理图案,然后建立物体表面的点与纹理空间的点之间的对应关系。另一类称为凸凹纹理,可用来表现外观凸凹不平的如未磨光的石材、从近处观察的有植被的景观等。生成凹凸纹理的方法是在光照模型计算中使用扰动法向量,直接计算出物体的粗糙表面。

综上所述,制作灰度立体图涉及的因素是比较复杂的。好在我们可以采取不少的简化措施,因为地表亮度的模拟不必十分精确,只要在不显著增加计算

量的前提下，有一定程度的真实感就行了。

灰度立体图的一个特例是正射灰度立体图。该类图采用垂直投影，相当于选择地表垂直上方无穷远处为视点；其光源通常指定斜射角度，使画面有较好的阴影感和立体感。正射灰度立体图相当于地图学中的地貌晕渲图，它在计算机中容易与矢量地图叠合，给矢量地图以立体背景。从高程矩阵中自动生成的正射灰度立体图与卫星影像和正射航片有时看上去类似，但是，它不包括任何地面覆盖信息，仅仅是数字化的地表起伏显示；它的光源可任意指定，一般确定为西北45°方向，而航片和卫星影像的光源是太阳。此外，它经过了平滑和综合处理，因而没有航片和卫星影像上显示出的丰富地形细节。

4. 进一步的三维表达技术

在灰度立体图的基础上，为了制作较高质量的三维表达，还有不少工作要做。这里介绍其中一部分。

首先，可在起伏地形表面叠加上真实的或仿真的地表或其地物。铺设仿真地面的材料可以是正射航片、卫星影像和线划地图等。图4-19即一幅地表铺上TM卫星遥感图像的浙江省分水江流域（局部）的灰度立体图图像。

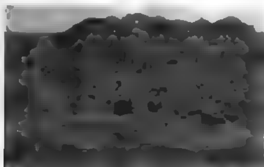
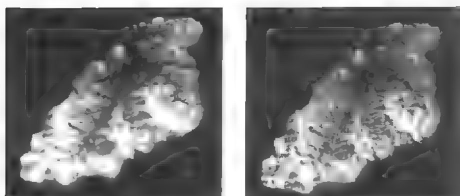


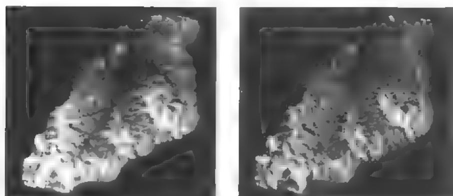
图4-19　在铺设TM图像的分水江流域地形上添加仿真地物

在有了仿真地表以后，还可以按照地图实际情况进一步添加仿真地物。例如，在有居民点的地方按地图建房或种树等。这些地物可通过额外的建模制成，并被植入到地表上。图4-19中的近景中和远景山脚下的房屋，就是按村庄所在添加的民房。

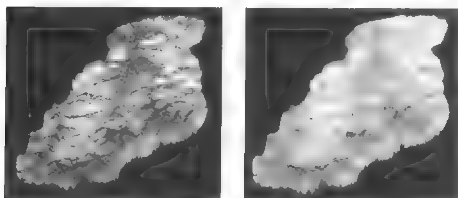
其他进一步的三维处理包括柔和度处理、高程夸张处理和环境处理等，如图4-20所示，该图表现浙江省湖州陆家庄小流域的地形。其中，图a表现柔和度处理，很明显，左图地形柔和，右图地形明显较左图尖锐，用户可根据需要调整地形起伏的尖锐度。图b表现高程夸张处理，这种处理用来夸大或减小地形起伏的程度。图c表现环境处理，即雾汽和水体等的处理。图中可见，环境



(a) 柔和度处理，右图地形较左图柔和



(b) 高程夸张处理 右图的高程起伏为左图的1倍



(c) 雾气控制 左图为无雾 右图为有雾

图4-20 三维地形表达的某些例子

有雾或没雾的形象有明显的差别。除4－20所示外，进一步三维处理还有光照控制、注记符号等，这里就不一一介绍了。三维动画和虚拟显示动态显示的内容，可参阅本书§7.2。

4.4.4 地形派生数据或地形分析

从数字高程模型可以导出一系列描述地形的，很有应用意义的派生数据，如坡度和坡向、等值线、视线图和各种地形轮廓线等。这些计算事实上就是进行地形分析。很多GIS软件已将这些处理手段固化进去，形成软件中的地形分析功能或模块。一部分GIS书籍也将这部分内容放在"空间分析"的章节中。

1. 坡度和坡向

在使用数字高程模型之前，地貌的描述和比较是用变化范围较大的定性或半定量分析技术，定量分析难以实现。GIS技术的发展使高程数据以数字形式产生高程矩阵或TIN系统后能用多种标准程序，进行坡度、坡向和其他地面特征的计算和制图工作。

数据指每个地面单元的坡度和坡向，其中坡度定义为地面单元与水平面之间的夹角，或地面单元的法矢量与铅垂线之间的夹角；有时也用这个角的正切（即斜率）来表示坡度。坡向为地面单元的法矢量在水平面内投影所指方位角，通常从北点方向顺时针度量。

计算坡度和坡向的方法主要有两种。一种方法是从三维空间位置矢量入手，计算地表单元的法矢量，掌握了法矢量就求得了坡度和坡向。在TIN数据情况下，三角形单元是一个平面，从三个顶点的(x, y, z)坐标容易确定三角面片的法矢量。在高程矩阵的情况下，每个地面单元，即网格的4个顶点严格说一般不共面，但我们仍然可以通过两对角线构成的空间矢量的叉积，来求取地面单元的法矢量。

在高程矩阵中，还可以利用活动窗口法，通过与周围网格高程的比较计算，通过简单差分的方法，快速地求取地面单元的坡度和坡向的近似值。这种计算常在3×3像元的窗口内进行，窗口在DEM数据矩阵中连续移动后完成整幅图的计算工作。每个像元的坡度和坡向确定后，就形成能可视化的坡度或坡向栅格图像。

地貌分析中有时还需要用到局部地表的凹率和凸率，这时可用包含足够像元的窗口进行相邻像元间的二阶差分计算。

2. 等值线

等值线也是一种从DEM导出重要地形派生数据。这种导出主要是测绘部门的工作，因为测绘部门同普通用户不同：普通用户通常从地形图上的等值线生成DEM；而测绘部门却常从航测相片之像对，直接快速地生成高程矩阵，或从

实测高程数据生成TIN，然后基于数字高程模型来生成等高线，制作地形图给广大用户应用。

从高程矩阵容易得到等高线图。事实上，按高程分层设置灰度的图4-17中，不同灰度带间的边界就是等高线。当然，这种等高线是锯齿形的，从制图观点来看还过于粗糙。一般而言，从高程矩阵绘制等值线须经过三个步骤：第一，计算各条等值线和网格边交点的坐标值，找到一系列等值点，这是一个数据插值的过程。第二，找出一条等值线的起始等值点（线头），并确定判断和识别条件，以追踪一条等值线的全部等值点。这是因为，相同高程的等值点连成的等值线常常不止一条；而且有的是闭合线，闭合必位于区域内部，其线上任一等值点均可作为线头和线尾；有的是开曲线，其起始等值点和终止等值点一定位于边界三角形的最外边上。需要具体情况具体处理。第三，联结各等值点绘制光滑曲线，可采用合适的样条函数等方法来实现这一点。

从不规则三角网DEM数据绘制等值线的步骤基本类似。首先在确定三角形边上存在等值点后，用内插法求得等值点的坐标。然后找出起始等值点，并追踪等值点。由于三角网比方格网稍微复杂些，线头找到后需要计数，且内插得到的等值点按三角形的序号排列，以便按一条等值线通过的先后顺序排列进行追踪。最后进行等值线的注记和曲线光滑。

3. 视线图

视线图亦称为可视域分析，是数字高程模型的又一种地形派生数据，它描述通视情况，即从指定观察地点视线能及的区域范围。可视域分析对军事活动、微波通讯网和旅游娱乐点的规划开发都有重要应用意义。数字高程模型（无论是高程矩阵或不规则三角网）的建立为这类分析提供了极为方便的基础，能方便地算出一个观察点或视点所能看到的各个部分。

视线图制作与上述消隐处理有些类似，但它不重视前景遮蔽物的细节，感兴趣的是可视域范围及其景观。具体制作时，从观察点所在的位置引出各方向的射线，比较射线通过的每个点（高程矩阵中即为像元）的高程，将不被物体遮藏的各点进行特殊编码，从而得到一幅简单的视线图。

在进行可视域分析时，一个需要注意的问题是DEM通常描述地面点的高程，而不包括地面物体，如森林和建筑物等的高度。因此，当地物高度对分析结果有不可忽略的影响时，需要考虑进行地物高度的因子修正，以正确地确定通视情况。

4. 地形轮廓线及其他

数字高程模型的地形派生数据还包括各种地形轮廓线，如山脊线、谷底线、集水线、流域范围等。数字高程模型上本身并没有这些线，但很容易用程序自动导出。

在高程矩阵中自动探测山脊线和谷底线，也可采用活动窗口和合适的算法，找出并标记相对高程最大或最小的一系列像元。连接它们就形成山脊线或山谷线。

集水线是山坡上高程等于给定数值的点的连线，集水线所包围的范围，即集水范围。集水线和集水范围可应用于水库等水利工程设计和洪涝灾害分析。在前一种情况下，可比手工方式精确得多地计算各种意义上的库容；在后一种应用中，可以近似计算某泛滥洪水水位所导致的淹没区，进而可以结合该区域有关数据进行洪涝灾害评估。

计算流域范围对流域分析十分必要。计算思路是：首先交互式地确定河流流域的出口，采用活动窗口搜索较"上游"的点。例如，采用 3 × 3 像元的窗口，将窗口中心像元置于起始点上，比较中心像元相邻近的 8 个像元的坡向。认为坡向最朝向中心像元的像元为"上游"，并以它为新的窗口的中心像元。重复比较过程，又能得到新的"上游"点，已有"上游"标志的不予比较。整个数据范围都运算完成后，流域范围就全部标记出来了。

DEM 数据还有其他用途，如线路勘察设计与土石方量估计等。土木建筑工程中的土石方计算与集水范围计算十分类似。事实上，集水范围是一种凹陷区的体积；而建筑工程中的土石方计算既可以是凹陷区，也可以是凸出区体积的计算。

第五章
空间分析

§5.1 空间分析一般概念与空间查询

5.1.1 空间分析概念

1. 关于“空间分析”概念的总体说明

GIS 学者普遍认为，空间分析是地理信息系统中最核心、最重要的概念之一，但各人对空间分析具体涵盖的理解不尽相同。本书中的空间分析，如§1.2 3.5 所述，是从比较广义的角度理解的，指 GIS 中一切用以提取地理空间信息乃至时空分布、组合、联系和发展的知识的、涉及空间位置要素和空间关系的分析。

空间分析在人类历史上有着悠久的历史与传统。事实上，人类出于生存和发展的需要，必须学会分析周围地理事物的空间关系，因而从一开始就始终在自觉或不自觉地进行着各种类型的空间分析。这种分析孕育了地理学。地图的出现使人类的空间分析能力大大加强，地图可视化和地图量测等手段导致空间分析水平的明显提高，人们得以利用地图进行较高层次上的空间分析，包括社会、经济、文化和军事等领域的各种区域性决策。在近代，传统的空间分析能力随着物理、数学概念与方法的引入和地学各分支的发展，又进一步大大加强。

现代“空间分析”概念的真实提出，源于 20 世纪 60 年代地理和区域科学的计量革命。开始阶段主要是应用统计分析的定量手段，分析点、线、面的空间分布模式；后来逐渐强调地理空间本身的特征、空间决策过程和复杂空间系统的时空演化过程分析；分析方法从统计方法扩展到运筹学、拓扑学和系统论等，定量分析能力不断加强。地理信息系统出现以后，出于其本质或天性，它不仅迅速地吸取所有能利用的空间分析方法和手段，将它们植入 GIS 软件中，而且应用各种计算机新技术，使繁难的传统空间分析任务变得简单易行，并能方便、高效地应用几何、逻辑、代数等运算、数理统计分析和其他数学物理方法，更科学、

高效地分析和解释地理特征间的相互关系及空间模式，从而使空间分析能力发生飞跃。

空间分析如此重要，以致在国内外所有的 GIS 书籍上，都有“空间分析”的章节，并给予重视和强调。然而，也许正是因为太重要，这个概念也缺乏一致明确的界定；众多的学者从不同角度、层面去理解、定义、使用，以致不同 GIS 书籍中“空间分析”章节所涵盖的范围、用词习惯，以及 GIS 软件开发商和用户对空间分析的理解和处理方式，皆不尽相同，有时甚至有较大差异。这种差异从根本上说，来自不同国家或不同部门所要求进行的，或所理解的空间分析的内容、深度和细度不同。有的国家或部门仅要求一般的空间查询或粗浅的分析，有的则将空间分析做得很深很细。作者认为，一个国家对空间分析的重视和需求的程度，是衡量其区域管理水平、社会发展水平和人的素质水平高低的重要标志。在我国，空间分析还远未能发挥应有的作用。

2. 本书的做法及空间分析概念框架

本书编写空间分析内容的思路是尽可能全面一些，其涵盖与邬伦等的《地理信息系统——原理、方法和应用》和龚健雅的《地理信息系统基础》有些类似，但本书主要讲解基本概念，不拟像他们那样详细和深入，主要目的是给学生一个较完整的“空间分析”知识框架的轮廓，知道空间分析可能包括哪些内容，每种分析及其方法是怎么一回事。作者从多年 GIS 教学实践经验感到，这样一种全盘性把握是学生所首先需要的。

为此，本书将从自己的视角对空间分析进行分类，并将仿照 § 4.1.1 的做法，构架一个较完整的“空间分析”内容和概念的框架。具体做法是：首先尽可能完全地“网罗”可能涵盖在“空间分析”大伞之下的内容或概念，形成一个“空间分析”内容或概念的集合。接着对上述集合中各种“空间分析”的内容或概念进行梳理，按有利于学习理解的原则分门别类，归纳成一个表达较完整“空间分析”的概念框架的一览表。然后，对该框架的思路和内容作一个总体说明，交代哪些内容将在本节顺便解释，哪些内容将放在本章后面哪一个章节用专门篇幅讲述。

这一番归纳整理的结果如表 5.1 所示。

3. 关于上述框架的说明

我们将空间分析分为空间查询、程式化空间分析和其他空间分析三大类。

首先，我们采取部分 GIS 学者的观点，将空间查询内容作为空间分析章节内容的一部分，尽管不少 GIS 书籍中的“空间分析”一章中并不提及空间查询。但是，我们将空间查询进一步分为基本空间查询和较复杂空间查询两类，它们分别在本节下面的两个分节中讲授。

表 5－1　空间分析概念一览表

<table>
<tr><td rowspan="6">空间查询</td><td rowspan="3">基本空间查询</td><td colspan="2">从地物空间位置（或图形）特征查属性</td><td>通过表示地物或区域位置的图形单元查询这些地物或区域的特征、内容或组成成分等属性</td></tr>
<tr><td colspan="2">从属性查空间位置（或图形）特征</td><td>通过地物或区域的属性数据或信息，查询满足给定条件的地物或区域的位置和形态，并加以可视化表达</td></tr>
<tr><td colspan="2">用户给定范围查询（拓扑空间区位查询）</td><td>查询用户给定的图形区域（矩形、圆和多边形等）内的地物属性和空间位置（或图形）特征，并加以显示，这种查询亦称为“query by graphics”</td></tr>
<tr><td rowspan="3">较复杂空间查询</td><td colspan="2">空间关系查询</td><td>邻接、包含、落入、穿越查询和缓冲区查询等</td></tr>
<tr><td colspan="2">维环境下的查询</td><td>维（动态或静态）环境下的位置与属性的交互查询</td></tr>
<tr><td colspan="2">与统计分析交互的空间查询</td><td>空间查询与空间现象相关关系分析之结合</td></tr>
<tr><td rowspan="4">程式化空间分析功能（已形成某些GIS软件中的固有功能）</td><td rowspan="2">空间操作</td><td colspan="2">传统GIS的空间操作</td><td>叠置操作、缓冲区操作、空间剪切、合并同属性多边形和合并处理中形成的细碎多边形等。这些操作通常同时涉及到空间特征（图形）数据和属性数据两者的较复杂处理</td></tr>
<tr><td colspan="2">OO空间操作</td><td>在面向对象数据模型基础上对空间对象进行空间操作</td></tr>
<tr><td colspan="3">网络分析</td><td>线状地物组成的网络系统中进行最佳路径、资源分配和地址匹配（选址）等分析</td></tr>
<tr><td colspan="3">维地形或环境的空间分析</td><td>在DEM和维显示技术基础上的进一步空间查询和空间分析，包括趋势面、表面积、体积、坡度、坡向、视亮度、集水线、脊谷和可视域分析等</td></tr>
<tr><td rowspan="3">其他空间分析</td><td colspan="3">空间统计分析</td><td>统计图表、分布密度和均值、主成分、层次、聚类、判别分析等，主要基于栅格数据形式</td></tr>
<tr><td colspan="3">分析地图学</td><td>数学地图学和地图分析，主要内容是基于矢量地图数据的统计学分析</td></tr>
<tr><td colspan="3">空间分析模型</td><td>应用领域专业模型与GIS输入、叠置或可视化相结合、空间区位配置信息模型、空间决策信息模型等</td></tr>
</table>

“程式化空间分析”是本书提出的分类，它泛指除空间查询外，业已形成GIS软件中之固有功能的那些空间分析手段，如空间操作、网络分析和地形分析或环境的空间分析等。之所以特别提出此分类是因为：第一，这些功能具有一定通用性质，对GIS应用具有普遍意义，学生很有必要了解这些功能的基本内容；第二，这部分空间分析已包含比较丰富的内容，业已具备成为一个大类别的条件。至于这个类别的名称，也可采用“GIS软件中的空间分析功能”或其他适

当的词。程式化空间分析将是本章第二节的内容。

空间分析的第三大类暂时称为"其他空间分析",指除上述两类外的所有空间分析。这些空间分析中不少属于提高层次,本书尽可能简明地介绍。这将是本章第三节的内容。

下面就来具体展开空间查询的内容。

5.1.2 基本空间查询

1. 关于"基本空间查询"的概念

空间查询检索是地理信息系统最基本的分析功能,能否有效地从地理信息系统海量数据库中检索出所需信息,将影响地理信息系统的进一步分析能力。尽管偏重空间查询但缺少其他空间分析功能的桌面绘图软件曾被区别于 GIS 软件(参见 §1.4.2.2 小节),尽管不少 GIS 书籍的空间分析章节中不包含空间查询的内容,但大多数 GIS 学者仍然把空间查询看做是空间分析的一部分,把桌面绘图软件看做是 GIS 软件的一部分。事实上,现代 GIS 的查询功能愈来愈强,大多数空间查询功能已不同程度地带有空间分析色彩,以致某些查询功能与空间分析已没有清楚的界限。这也是作者在本章介绍空间查询的一个原因。

在现代地理信息系统的空间查询中,有一部分空间查询功能具有重要的意义,这就是几乎所有 GIS 软件或桌面绘图软件都具有的空间查询功能;如果连这些空间查询功能都没有,就谈不上是 GIS 软件。换言之,这部分空间查询功能是 GIS 软件的最基本的标志。正因此,我们将它们称为"基本空间查询",而将其他的空间查询功能称为"较复杂的空间查询"。

本小节所要讲述的内容就是基本空间查询。

一般而言,较复杂的空间查询比基本空间查询的分析色彩浓一些。但即令如此,基本空间查询的分析色彩仍然是不可忽视的。下面将看到,很多查询的条件和结果本身都有明显的分析意义。另一方面,将查询检索的结果用图形图像等可视化方式表达出来,也是空间查询检索功能的必要组成部分;而可视化表达对辅助用户进行分析有极大的意义。

还应当说明,从广义上讲,基本空间查询功能应当包括单纯在地理属性之间的查询,如查询地名为"宁波"的城市 2001 年的人口和总产值为多少等;基本空间查询也应当包括查询点、线、面的最基本的几何数据,如坐标、长度或面积等。但是,正如我们在 §4.1.2 中所指出的,地理属性之间的互查询是一般数据库系统的功能;基本几何数据的查询是一般图形处理软件也有的功能。它们都不属于 GIS 软件"基本标志"的范畴。GIS 特色的基本空间查询,必须涉及空间图形(位置)和属性的交互查询。下面我们具体介绍三种类型的 GIS 基本空间查询功能。

2. 从地物空间位置(或图形)特征查询属性

第一种基本空间查询功能是通过表示地物或区域位置的图形单元(点、线和多边形)或图像单元(像元),检索查询这些地点、地物或区域的特征、内容或组成成分等属性。这其实就是§1.4.1中列出的地理信息系统五项主要任务中的第一条:"某个地方有什么(属性)"的问题,英语中有时称为query by location。在矢量GIS系统中,这种空间查询可以简记为"图形⇒属性查询"。图形⇒属性查询的最常用方式,是利用鼠标直接点击所要查询的图形。GIS系统给出查询结果的方式,是多种多样的,可以采取文字、列表方式,但更多是以可视化方式,例如在屏幕或输出地图上,以鲜异的色彩或高亮度显示满足给定条件的特征。

在图形⇒属性查询中,最普通的就是基本几何参数的查询,一般GIS软件都提供这种默认的功能,虽然它不是GIS的特色。

实现图形⇒属性查询需要通过属性-图形的标识符,在关系数据库中检索到该图形所对应的属性,并显示出来。

3. 从属性查询空间位置(或图形)特征

这是前一种查询的相反过程,就是通过地物或区域的属性数据或信息,查询满足给定条件的地物或区域的位置和形态,并加以可视化表达。为方便起见,下文中将此类查询简记为"属性⇒图形查询"。这种查询有时称为query by attribute,它实际上就是§1.4.1中列出的地理信息系统五项主要任务中的第二条"符合某些条件的实体在哪里"的问题。由此也可见,我们所说的"基本空间查询"功能包含了地理信息系统五项主要任务中的两项,确实具有GIS"基本标志"的意义。

在属性⇒图形查询中,属性条件可以采取单纯属性、组合属性和模糊属性等方式。单纯属性的查询方式指查询某属性为某值的图形或地物,这是最简单的形式。组合属性方式指给出多种属性的一个组合,例如,查询某省内,面积超过1 000 km^2、人口大于30万的城镇有哪些?库容量和发电量分别大于某值,且大坝建成时间超过10年的水库在哪里?由此可见,组合属性条件本身就带有较浓的分析色彩,为用户展开了更大的查询分析空间。模糊属性方式指给出的属性条件是不够明确的,或不完整的,这样的查询称为模糊查询。模糊查询也是带有分析色彩的重要查询方式,并有重要的应用意义,因为它能帮助用户实现本来难以完成的查询。例如在土地管理GIS系统中,欲查询一宗地,只记得该权利人姓周,所在村庄和宗地号部分数字,也可能通过这三个模糊属性条件,查询到所要的结果。

一种有特别意义的属性⇒图形查询是地址匹配查询,这是根据街道的地址,包括地理编码和街道的门牌号码等,来查询事物的空间位置、所在街区和其

他属性信息。地址匹配是地理信息系统特有的一种查询功能，它对空间分布的社会、经济调查和统计很有帮助，只要在调查表中添了地址，地理信息系统可以自动地从空间位置的角度来统计分析各种经济社会调查资料。这种查询也经常用于公用事业管理和事故分析等方面，如邮政、通讯、供水、供电、治安、消防和医疗等领域。由于地址匹配的分析意义，有些 GIS 学者将它归于空间分析范畴。由此也可见，空间查询和空间分析之间的交混关系。

实现属性⇒图形查询自然也需要通过的属性 - 图形的标识符和数据库，但单纯属性方式的实现比较简单。这种查询仅需采用简单语句，找到满足条件的数据库记录，得到它的目标标识，再通过目标标识在图形数据文件中找到对应的空间对象。而组合属性和模糊属性等方式则需要构架相应的 SQL 语句，检索到满足组合条件的数据库记录，再通过目标标识在图形数据文件中查找。SQL 语言是关系数据库技术集数据操纵、数据定义和数据控制为一体的 DBMS 标准数据库语言（见 § 3.1.3.2）。

不过，构架 SQL 语句的工作并不要求用户去做。一般的 GIS 软件都设计了比较好的用户界面，例如给予用户一个输入属性条件的对话框，用户通过对话框来直接给定属性条件。例如，在图 1 - 6 所示对话框中，用户在“产权证”一栏中输入证件号，就可以查询到该宗地位置和其他属性。很多 GIS 系统的对话框中把用户常用的属性列为默认的可选项，让用户用鼠标去选取，而免去输入的麻烦；在需要的地方，对话框还允许用户填写模糊属性、逻辑表达式或数学表达式等，来指定一个或一组条件，用户给定的条件实际上就是 SQL 查询语句中要求填写的条件，GIS 系统会自动将用户填写的内容植入 SQL 语句中去。

4．用户给定的几何窗口查询

第三种基本空间查询是用户给定一个图形范围，如矩形、圆和多边形等窗口，查询这个范围内某类地物的所有对象，包括其属性和位置（或图形）双重特征，并加以显示，必要时还要显示出每个对象的属性表。这种查询亦称为空间定位查询、窗口查询或图形查询（query by graphics）。

定位查询的窗口形状通常有矩形、圆形和多边形等。矩形查询是给定一个矩形窗口，查询出该窗口内某一类地物的所有对象。按圆查询是给定一个圆或椭圆，检索出该个圆或椭圆范围内的某个类或某一层的空间对象。按多边形查询，指让用户用鼠标给定一个多边形，或者在图上选定一个多边形对象，检索出位于该多边形内的某一类或某一层的空间地物。多边形查询非常有用，因为经常需要查询某一区域，例如，查询行政区所包含或涉及的某类地物，如穿过的公路或所包含的大型水库等。这时，用户可以用鼠标点取该区域的边界线（一个多边形）进行查询。

窗口查询给予用户方便而有力的查询工具，但开发窗口查询功能比较复

杂。首先需要根据空间索引，检索到哪些空间对象可能位于该窗口内，然后根据点在窗口内，线在窗口内，多边形位于窗口内的判别计算，检索出所有落入检索窗口内的目标。特别是多边形查询，涉及到点在多边形内、线在多边形内、多边形在多边形内的判别计，尤其复杂。窗口查询往往还需要处理两种选择：一种是只检索出包含在该窗口内的地物；另一种是检索出窗口涉及到的所有地物，包括全在窗口内的地物和部分位于窗口外的地物。例如在旧城改造区规划一条新的道路，涉及到很多房屋拆迁。显然，拆迁的房屋不仅应包含新道路(长条形矩形窗口)内的房屋，还要包括所有与新道路边界线相交的，即路内、路外各占一部分的房屋。

5.1.3 较复杂的空间查询

1. 空间关系查询

空间关系查询是一种基于复合条件的查询。与上述属性组合查询不同，这里的空间关系查询不仅涉及属性，而且涉及空间几何关系，包括空间拓扑关系的处理。上述几何窗口查询严格说也是涉及空间几何关系的查询，但是，它们与本小节所说的空间关系查询有区别。几何窗口查询是一般 GIS 软件中都有的功能；而本小节的空间关系查询不然，一般不是 GIS 软件必备的功能，因而不属于基本空间查询。

本小节的空间关系查询虽然不是 GIS 软件必备的功能，但它们却是 GIS 的常见功能，事实上，大多数 GIS 软件，包括桌面绘图软件，都或多或少具有空间关系查询的功能，只是功能强弱不同而已。当然，拓扑空间关系查询除外，因为一些没有拓扑结构的 GIS 桌面绘图软件，如 MapInfo，不可能具有空间拓扑关系的查询功能。

空间关系查询有两种主要方式。

第一种是直接通过拓扑数据结构实现的空间拓扑关系查询。当然，这种方式只适用于具有拓扑结构的数据。例如，在具有拓扑属性表的 GIS 软件中，查询与给定多边形邻接的多边形，可通过拓扑查询解决：首先从弧段与多边形关联表中，检索出组成该多边形的所有弧段或线。例如在表 2－3 中，查找组成 2 号多边形的弧段可用检索式 Lpoly = 2 或 Rpoly = 2。接着，再从弧段关联表中，检索出这些弧段所在的另一侧(左或右)的多边形。

基于表 2－3 这样的拓扑关系属性表，还可以进行线状地物联结性查询，例如查询某河流的所有支流；可以进行线、点查询，例如某条道路上有哪些桥梁，某条输电线上有哪些变电站等；还可以进行面－线查询，例如某个多边形的边界有哪些线等。此外，多边形对点、线和较小多边形的包含关系的查询，有时也可以通过拓扑关系数据来实现。

然而，有拓扑关系可利用的场合毕竟有限，在更多的情况下，空间关系查询需要通过空间几何关系运算来实现。例如落入查询，即查询一个点、线或面状地物象位于哪个多边形之内，就需要进行“点在多边形内”、“线在多边形内”或“面在多边形内”的空间运算和判别。包含关系查询，即查询某一个面状地物所包含的点状地物、线状地物或面状地物，也要进行类似的空间运算判别。又如，穿越查询，即查询线状地物，如某条公路或河流穿越了哪些省、县或乡镇等，需要进行线相交计算，即根据所查线状地物的空间坐标，计算出它与哪些面状地物或线状地物相交，有交点者即为查询结果。

有两种比较重要的空间关系查询需要特别提及。

一种是缓冲区查询。缓冲区（buffer）是点、线或面状地物的周围邻域的统称，其具体定义是：与点、线或多边形边界的距离为给定值的区域。例如，下一节的图 5－4 中就有一个距绍兴夏履江 200 m 的缓冲区。缓冲区查询是一种典型的近邻查询或邻域查询，在资源、环境等区域分析中有广泛应用。例如，查询某地区内，距离某种污染源 100 m 以外、满足若干地质条件的地下水源最可能在那里？京沪线之东、距离京沪线不超过 50 km，且城市人口大于 100 万的城市有哪些？

缓冲区查询与本章下一节将较详细讲述的“缓冲区分析”概念非常接近，请读者参阅该处的内容。这里仅说明两个概念的差别：缓冲区查询与上述多边形查询类似，完全以查询为目的，因而它所生成的缓冲区是暂时的，只作为近邻查询的辅助线，查询完后缓冲区不保留在原图上，且不一定要求是拓扑数据；而下一节中的缓冲区分析重在分析结果，常采用拓扑数据，生成的缓冲区常作为新的多边形文件保存。

还有一种较重要的空间关系查询是线路查询。线路查询是在线路网络，如交通道路网、旅游网、河网和各种城市管道网等网络中所进行的查询，它与下一节中的“网络分析”也非常接近；差别只在于：网络分析旨在分析出最佳结果，而线路查询仅给用户以查询结果，让用户自己去判断和选择。

以旅游或交通网络为例。有一种“等时查询”要求查询指定时间内，从给定中心点出发旅游所能达到范围内的全部景点，并给出各景点的文字说明或图像显示，供用户选择。又如选线查询，在这种查询中，用户可选择各种旅游线路，查询沿线各旅游景点及服务设施；或选定不同的旅游始点和终点，查询两点间的各种旅游线路方案，并进行选择等。近年来，国内出现了很不错的地图网站，网民可从中进行国内各大城市公交网查询，如查询给定始点和终点站间的若干种公交乘车、换车的方案。

空间关系查询的种类还不止上述几种。但从中我们已可感到，开发软件中的空间关系查询功能会涉及到更多的技术问题，它需要将 SQL 查询与空间图形

查询结合起来。为此,不少 GIS 学者提出将 SQL 的属性条件和空间关系的图形条件组合在一起,形成面向空间数据的扩展 SQL 查询语言,或更进一步的空间查询语言(spatial query language)。这是一个很有意义的 GIS 研究课题。解决该问题可采取两种思路,一是将属性条件和空间关系整体统一起来,从底层进行查询优化;二是将两层分开进行查询优化。前一种思路解决问题更彻底,但技术难度较大,仍处于理论发展和技术探索阶段。

2. 与统计分析交互的空间查询

另一种较复杂的空间查询是与统计分析交互的空间查询。典型的例子就是 § 1.4.1 4 中引述过的酒吧攻击事件分析,该例取自 1999 年 2 月美国国家科学基金会和美国大学地理信息科学大学联合会共同举办的研讨会文献"地理信息科学:一个新兴的学科交叉研究领域中的紧迫问题"(David M. Mark 编辑),该文献想象了 2010 年 GIS 在三种领域的应用状况。其中,社会学领域的应用就是分析酒吧攻击事件。这里,我们将该段部分文字引述如下:

一个社会学家正在研究美国东北部一城市的犯罪现象,试图发现攻击事件的空间结构特征(pattern)。坐在多媒体系统前,他要求在城市地图上显示过去年份的所有攻击。在观察地图的同时,GIS 系统自动计算与现有其他数据的相关关系,并且注意到几个与犯罪数据有空间关联的现象。其中之一是攻击事件分布与酒吧密度之间的相关。研究者接受了这一特别建议,GIS 系统就在地图上添加了酒吧的位置。接着,研究者打开了一个建模窗口并定义了一个规则:"酒吧攻击"是当地时间下午 5 点至夜间 2 点之间、距离任一酒吧 100 m 范围内的攻击。然后他要求用黄色点显示酒吧攻击,用红点表示非酒吧攻击。接着,他让 GIS 系统显示区域照明情况、交通格局和警察巡逻方式。GIS 系统自动地进行空间统计分析计算,并给出计算结果……

从这个例子我们可以清楚地看到一种空间查询和空间统计分析相结合的工作模式:对可能相关的空间变量进行空间统计分析和空间查询;将分析和查询的结果可视化;一旦发现某些空间变量分布之间所存在相关关系或空间结构特征(patterns,例如酒吧攻击),就针对这些空间结构特征与其他空间变量(区域照明情况、交通格局和警察巡逻方式等)的关联做进一步空间统计分析和查询……

由此可见,与空间统计分析交织的空间查询比基本空间查询带有更浓的分析色彩。

3. 虚拟现实环境下的空间查询

虚拟现实环境下的空间查询是近几年来才出现的空间查询功能。由于技术实现的难度,虚拟现实环境下的空间查询的内容还不多,目前只限于图形⇒属性查询和属性⇒图形两种。但是,它具有很大的应用意义和前景,因为三维

虚拟现实环境的GIS具有更高的空间表达和可视化能力。

从原理上说，虚拟现实环境下的空间查询主要将视平面上的地物位置，通过透视和投影关系，与该地物原有的地理位置坐标联系起来（§4.4.3），并通过原来地物坐标与其属性数据之间的标识符，链接到其属性数据。在确定地物原有的地理位置坐标时，需要考虑到视平面上的点与原空间的物点的关系可能是一对多的关系，因为原空间近处和远处的许多地物可能投影到视平面的同一位置，但视平面上可见的地物肯定是未被消隐的地物。因此，消隐处理时最好将被消隐的和未被消隐的地物数据分开，单独存放。

尽管如此，由于三维-虚拟现实表达所涉及的巨大数据量和其他种种原因，实现虚拟现实环境下的空间查询还是不容易的。

图5-1是虚拟现实环境下的空间查询之一例。



图5-1 浙江省湖州市戴家山流域虚拟现实环境中的空间查询

§5.2 程式化空间分析

GIS空间分析技术经历了几十年的长足发展。在大量空间分析的研究和实践中，一些常用的、对不少应用领域都有普遍意义的空间分析手段被总结、提炼出来，形成通用型GIS商业软件包中的一些固化的空间分析功能或模块，如各种空间操作模块、网络分析模块、三维地形分析模块和空间统计分析模块等。本章的"程式化空间分析"，就是这些业已在GIS软件中形成固有功能的空间分析手段的汇编。

并非所有的GIS软件都有这样的功能；即使拥有，拥有的多少和强弱也不同。一些偏重空间查询和制图的桌面绘图软件就较少拥有这些程式化的分析

功能。程式化分析功能业已成为GIS软件空间分析能力的一种标志，GIS业界在评论具体GIS软件分析能力强弱时，常看该GIS软件所包含的程式化空间分析功能的多少及强弱。在国际GIS软件中，具有较强空间分析能力的代表是美国的Arc Info；在国内GIS软件中，空间分析能力较强的是GeoStar。

但应当说明，虽然我们定义的程式化空间分析包括地形分析和空间统计分析，但本章未包括这两项内容。这是因为，地形分析已在§4.4.4讲解；在空间统计分析方面，虽然有些GIS软件，特别是遥感图像处理软件形成了一些固化的空间统计分析模块或功能，但是，其中大多数是通用的统计学手段。因此，与其说将空间统计分析作为程式化空间分析的内容，还不如从更广泛、更基本的角度来考察空间变量的统计分析方法。这就是我们将空间统计分析放到下一节讨论的原因。由此缘故，本节的内容主要是空间操作和网络分析。

5.2.1 空间操作及分析

1. 空间操作的概念

本书所说的"空间操作"指Arc/Info为代表的GIS软件所形成的叠置分析、缓冲区分析、空间筛选、空间剪裁和多边形合并等程式化的空间分析功能。这些功能具有鲜明的GIS特色，因为它们都要同时处理空间特征(图形)数据、属性数据以及二者的挂联关系，且常常基于拓扑关系数据。例如，当筛选或剪裁下一块地图时，总是连图形带属性地被筛选或剪裁下来，被筛选或剪裁下来的图形仍然是带着各自的属性；必要时，GIS软件还要为被剪裁或修改的地图重新建立拓扑关系。

这些空间操作都是空间分析，而且都是比较"经典的"空间分析。我们采用空间操作一语，第一是因为，Arc/Info及其用户在有关文献曾将这一套"经典的"空间分析称为空间操作(spatial operation)；第二，需要有一个用语统称这一套空间分析功能，以区别于其他的空间分析；第三，即使是这一套空间分析，有时也有层次之分。这时，我们将这些空间分析软件功能的直接实施称为空间操作，有利于区分不同的分析层次。例如，矢量GIS的叠置分析有层次之分，第一层是对多层专题图直接实施GIS软件的叠置分析功能(重新找交点，分配属性，重新建立拓扑关系，并形成包含各专题因子属性的新图层等，见§4.1.2.2)；第二层叠置分析是在此基础上的进一步多因子综合分析。不少GIS书将这两个层次皆称为叠置分析，不利于初学者分清这两个层次。为此，我们可将叠置分析中的第一层次专称为叠置操作(详见本节后文§5.2.2.3)，以区分进一层的叠置分析。

下面介绍各种具体的空间操作及分析。其中，叠置分析与缓冲区分析内容较多，我们将用专门的篇幅讲授。此外，由于近年来空间操作一词又有了新的

含义,指基于面向对象数据模型的、对空间对象所进行的操作。因此,下面的内容将不仅包括“传统”的空间操作,也简介新意义上的空间操作概念。

2. 空间剪裁与空间筛选

空间剪裁与空间筛选都是从已有图层中抽取或取代一部分;但是,空间剪裁是按照用户给出的图形窗口或地图进行剪裁,而空间筛选是按照用户指定的属性条件,通常是组合的属性条件来进行筛选。不论哪种情形,操作都是连图形带属性进行的,即结果都是完好的 GIS 图层,在拓扑数据条件下,这些操作还可能自动为结果层建立拓扑关系。

空间筛选允许用户通过属性表从范围和内容两种角度进行筛选。例如,假定有一幅浙江省地图,包含地区(或专区或地级市)和县(或县级市)行政区界线、道路、河流和居民点等,我们可以利用线属性表从中选取其中部分内容,如纯粹行政区界线,或道路加居民点的浙江省地图;也可以多边形属性表从中选取杭州市部分的地图(因为杭州所辖区域都有“杭州”属性)。如果需要,还可以从杭州地图中再选出部分范围或部分内容的地图。

空间剪裁与上一节的窗口查询(或多边形查询)有一定相似性。但是,第一,这里的剪裁窗口或框线,往往是现有地图的边界或区域边界线;第二,这里不以查询为单纯目的,它要求剪裁的结果并加以保存,为进一步空间分析服务;第三,这里还不仅进行窗框剪裁,还可进行局部的地图替代。空间剪裁操作包括下列四种:

一种为“分拆”(split),即把一个图层按用户自画线条或指定的原图内线条,剪分为若干部分。例如,利用地区界线将浙江省地图剪分为杭州、宁波、绍兴和温州等地区(专区)地图。

另外三种如图5-2所示,图中左边平行四边形表示原图层;中间平行四边形中的椭圆表示用以剪裁原图层的框线或地图,这种地图通常为一幅区域地图,地图边界就是区域界线;右边平行四边形表示操作结果;“+”和“→”表示二者进行剪裁操作并得到结果。图5-2所描述的三种空间剪裁操作分别是:

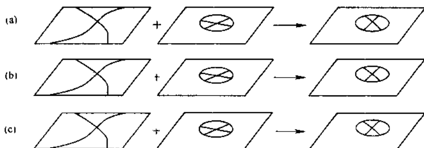


图5-2 空间剪切操作示意图

一种为"剪下"(clip),即按指定框线或另一幅地图的边界,从一个图层中剪下一块地图。如图5-2a所示,用椭圆图框剪裁的结果为椭圆框线圈出的原图层局部。

一种为"抹去"(erase),即按指定框线或另一幅地图的边界,从一个图层中抹去一块地图,如图5-2b所示。

另一种是"更新(或增改)"(update),即用另一幅地图去取代一个图层中的一部分,如图5-2c所示。

空间剪裁与空间筛选为空间分析和制图提供了非常灵活的手段和广阔的空间,一项分析或制图任务,可以通过多条途径达到目标。例如,欲从浙江地图中得到温州市河流道路图,既可以先从省图中剪下温州市地图后,在通过属性选择出温州市河流道路图;也可以通过属性先从浙江省图中选择出浙江省河流道路图,再剪下温州市河流道路图,等等。

3. 多边形合并

多边形合并包括两种情形。一种是合并处理中形成的细碎多边形。细碎多边形在空间数据处理中经常发生。例如,同一区域不同来源的两幅地图不可能完全相同(参见图2-8),其线状地物和多边形边界常常是非常接近但不重合,叠加在一起就会有很多沿线的细碎多边形。常发生细碎多边形的情况还包括:不同的人或同一人在不同时期数字化同一地图,以及基于Spaghetti数据结构(回顾§2.2.2.1)的软件输入数据时,每个多边形边界要数字化两次等。正因此,自动合并细碎多边形是GIS软件中十分普遍的功能。当然,细碎多边形通常不可能百分之百地自动合并,少部分合并还需要人工干预。注意,这里我们说合并细碎多边形,而未用"消除"一词,是因为消除细碎多边形,就是摘掉其一半的细长边界,这实际上就是将多边形合并到两侧的某个大多边形中去。

另一种多边形合并是相同属性的相邻多边形之合并。在有拓扑结构的数据中,这种操作容易实现,因为一条线的左多边形和右多边形是否具有某种相同的属性,是容易检索的;检索到这样的两个多边形后,去掉它们中间的公共边,就意味着二者合并。在Arc/Info等GIS软件中,有合并相同属性的相邻多边形的现成功能。

合并相同属性相邻多边形的操作可用于制图综合中的质量综合。例如,1:25至50万比例尺的土地利用现状图可以只采用国家2级分类(回顾§2.1.4.2和表2-1),在将较大比例尺土地利用图缩编为1:25万时,原有的3级地类就应合并为2级分类。例如,编码分别为311、312、313、314、315和316的用材林、防护林、经济林、竹林、薪炭林和特殊用林,应当合并为编码为31的"有林地"。对此,我们只需通过简单的指令,在多边形属性表中,将地类编码为"31×"的属性全部改为31,然后,直接调用相同属性相邻多边形合并的软件功

能，将地类编码为31的相邻多边形全部合并，就完成了有林地的质量综合。

4. 在面向对象的数据模型基础上对空间对象的空间操作

近年来，发展面向对象的空间数据模型（参见 §3.1.4.2）以克服传统GIS的不足，是GIS重要的发展方向。按照面向对象的空间数据模型，地理空间被看做是许多空间对象（城市、村庄、道路和河流等）的集合，这些对象都具有自己的属性。定义一个空间对象需要定义它的状态和行为。对象行为由一些操作定义，这些操作作用于一个或多个对象（运算对象），并产生一个新的对象（结果）。我们可以先定义一系列基本对象（如点、线和面），然后通过一系列操作构建应用域中所有其他对象。这就是新意义上的空间操作。作用于空间对象的空间操作可分为两类：静态的和动态的。静态操作不会导致运算对象发生本质的改变，如计算弧的长度并没有改变弧本身，动态操作会改变（甚至生成或删除）一个或多个运算对象，例如创建、删除和更新这三个基本的动态操作，以及再制作、生成、分裂、合并、转换、更新和缩放等。进一步讨论超出本书的范围，这里引述了少许内容，是为了使读者对新意义上的空间操作有一点感性观念。有进一步兴趣者可参阅陈述彭、鲁学军、周成虎的《地理信息系统导论》的第四章第三节。

5.2.2 叠置操作及分析

叠置（overlay，也译为叠加或叠合）操作及分析，主要用于多重专题图层的综合分析（见§1.2.3.2、1.3.1.2和1.4.2.2），是GIS的区域性和多层次特点所决定的传统空间分析手段。叠置操作及分析可分为矢量GIS叠置和栅格GIS叠置两种主要类别；而矢量型叠置又包括点与多边形叠加、线与多边形叠加和多边形叠加三种主要形式。

1. 点与多边形叠置

点与多边形叠置，指一个点状地物图层与一个多边形图层相叠，叠置结果通常是将其中一个图层的属性信息"注入"另一个图层中，更新该数据层面；然后基于新图层，通过属性查找获得点与多边形叠加所需要的信息。

实现点与多边形叠置实际上有两个步骤。第一步，计算多边形对点的包含关系。矢量结构的GIS能够通过计算每个点相对于多边形线段的位置，进行点是否在一个多边形中的空间关系判断。

第二步，进行属性数据处理。比较简单的方式是将多边形图层的属性数据，注入到点数据层的属性表中，使点属性表中含有每个点所在的多边形标识序号，以及多边形的某些附加属性。例如将油井（点）图层与行政区划（多边形）图层叠置，油井的点属性表将更新，该表不仅包含原有的油井本身属性，如井位、井深、出油量等，还添加了所在行政区的标识序号，以及行政区名称和相关经济数据等属性。当然，也可以将点的属性注入到多边形上，更新多边形属性表。但是，这样做比较麻烦，因为多边形可能包含不同数目的点，在多边形属性表中，可能需要添加若干个点的标识和属性。由此缘故，点与多边形叠置一

般都是将多边形属性记入点数据层的属性表中，而不是相反。

通过点与多边形叠置可以查询每个多边形里有多少个点，以及落入各多边形内部的点的属性信息。例如，一个全国矿产分布图（点）与一个中国政区图（多边形）经叠置分析后，通过更新的点属性表，可以计算各省有多少种矿产，产量有多少；或者查询指定类型的矿产在哪些省里有分布等信息

2. 线与多边形叠置

线与多边形的叠置，指一个线状地物图层与一个多边形图层相叠，叠置结果通常是将多边形层的属性注入到线状地物层中，更新其线属性表；然后在线数据层基于更新的属性表，查询分析叠置信息。

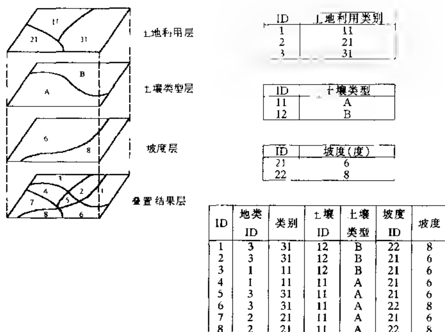
实现线与多边形叠置也分为两步。首先要比较线坐标与多边形坐标的关系，判断哪一条线落在哪一个或哪些个多边形内。后者是因为一条线常跨越多个多边形。为此，必须计算线与多边形边界的交点；以交点作为结点（参见§2.2.2.2 结点概念），将原线截为2个或以上的、落入不同多边形的新弧段。第二步是重建线的属性表，表中既包含每条新弧段原来所属的线的所有属性（如原线为某种道路，被截成的新弧段仍有该类道路的属性），也包含新添加的、它所落入的多边形标识序号，以及该多边形的某些附加属性。

这样，根据新的线属性表，就可以确定每条弧段落在哪个多边形内，多边形内的新弧段以及多边形其他信息。例如，将河流图层与省市行政区多边形图层相叠置，河流图层中每个新河流段的线属性表不仅含有原河流的信息，还含有该流段所在省市行政区的标号或其他信息。据此，就可以查询任意省市内的河流长度，计算它的河流密度，分析河流与所在省市其他属性的关联等。如果线状图层为道路网，叠加的结果可以得到每个省市内的道路网密度，内部的交通流量，进入和离开某个省市的交通量，相邻省市之间的相互交通量等。

3. 多边形叠置

多边形叠置是叠置分析中最经典的形式，也是 GIS 最常用的空间分析功能之一。多边形叠置将两个或多个多边形图层进行叠加，产生一个新的多边形图层。新图层的多边形是原来各图层多边形交割的结果，每个多边形的属性含有原图层多边形的所有属性数据，以及在原图层中的序号（ID），如图 5－3 所示。图中给出土地适宜性评价常用的三个因子层：土地利用现状、土壤类型和坡度。图中可见，叠置结果层每个多边形含有原来三种属性数据，从中可以看出该多边形与原图层的对应关系。例如，结果层1号多边形在土地利用层中属于属性为31的多边形，在土壤层中属B类，在坡度层中位于8°地带。

多边形叠置操作的技术实现也分为两个步骤：第一步为几何求交，即首先求出所有多边形边界线的交点，再根据这些交点重新装配多边形，建立拓扑关系，每个多边形赋予唯一标识码，并判断新生的多边形分别落在每个参与叠置



ID	土地利用类别
1	11
2	21
3	31

ID	土壤类型
11	A
12	B

ID	坡度(度)
21	6
22	8

ID	地类ID	类别	土壤ID	土壤类型	坡度ID	坡度
1	3	31	12	B	22	8
2	3	31	12	B	21	6
3	1	11	12	B	21	6
4	1	11	11	A	21	6
5	3	31	11	A	21	6
6	3	31	11	A	22	8
7	2	21	11	A	21	6
8	2	21	11	A	22	8

图5-3　多边形叠置表示之一例

的多边形层的哪个多边形内，建立新生多边形与原多边形的关系；第二步是在关系数据库中建立结果层的多边形属性表，将原图层中对应多边形的属性数据，拷贝进（或关联）到新的多边形属性表中。这种属性分配方法的理论假设是多边形对象内属性是均质的，将它们分割后，属性不变。

多边形叠置操作还有一个需要处理的技术问题是结果层的边界确定问题。在图5-3中，参与叠合的各层地图范围一致（例如均为标准图幅），结果层边界自然为原图幅。但是在实践中，参与叠合的各层地图边界不一定一致。因此，GIS软件通常要让用户决定结果层的边界选取。例如，在两个图层叠置的情况下，通常有这样几种选择：或者取其中一层地图的边界，或者取二层地图范围之交集为结果层的范围。总起来说，多边形叠置操作的技术实现比前面两种叠置要复杂。

多边形叠置操作完成后，可以像前面两种叠置一样，基于新图层的属性表综合查询参与叠合的多因子属性信息；多边形叠置操作也可以用来进行缓冲区分析（见下一小节）等。然而，应当特别指出，人们使用多边形叠置操作，一个经常的目的是为了在此基础上做区域多因子综合分析，特别是进行各种资源与环

境领域的分析评价。再以图 5－3 为例，用户在进行了如图所示的多边形叠置操作之后，可能在其基础上，采用某种分析运算关系 f 来进行某种土地评价：

$$分析评价结果\ E = f(土地利用因子、土壤因子、坡度因子) \qquad (5-1)$$

为此，他需要在图 5－3 中结果层的多边形属性表中增加表示分析评价结果 E 的字段，用以存储结果层中每个多边形的分析评价结果数据。再后，他还会利用 GIS 平台显示叠置结果层，可视化地表达、观察和分析评价结果 E 的空间分布。

由此可见，在多边形叠置中，确实存在两个层次的工作，一个是多边形叠置操作；另一个是建立在叠置操作基础上的进一步的空间分析，或多因子综合分析评价的工作。

从上面还可看到，多边形叠置操作结果层上的多边形，就是区域性多因子综合分析评价的基本单元（有时也称为最小公共地理单元）。由于新图层的多边形是原来各图层多边形交割的结果，它们总是比原图层的多边形碎小，且不均；特别当参与叠加的因子图层数目较多，或原图层本身的多边形已较密时，结果层的多边形将过于细密和不均。这些细密多边形并不是前述需要合并消除的、数据处理过程中产生的不必要多边形，而是有其实际地理意义的分析单元；但是，地域单元过于细密毕竟会带来合理性、可视性等方面问题。这是矢量 GIS 的多边形叠置的一个欠缺。

4. 栅格 GIS 的叠置分析

栅格 GIS 的叠置分析概念我们已在 §2.3.2.2 中讲过。从应用功能上讲，栅格 GIS 叠置分析与矢量 GIS 的多边形叠置类似，常用于资源与环境领域的分析评价。事实上，20 世纪 60 年代早期加拿大国家土地调查局开创的世界上第一个地理信息系统，进行的就是栅格数据形式的空间叠置分析（见 §1.3.1.2）。同矢量多边形叠置相比，除去矢量和栅格两种数据结构本身的特点不谈，栅格 GIS 的叠置分析有两个明显的特点：

第一，栅格 GIS 的叠置分析中各图层格网一致（参见图 2－11），叠置过程中分析评价的基本单元，或最小公共地理单元不变，因而不存在矢量多边形叠置中结果层基本单元可能过于细密的问题。早先由于计算机能力的限制，栅格 GIS 的网格有时不得不粗一些，现在这一点也不成问题了。因此，在现代分析评价中，当因子很多时（十几种甚至几十种），多采用栅格型叠置分析，简单而有效。当然，从技术上说，如果栅格 GIS 的叠置采用的是压缩格式的栅格数据，还需多做些工作。

第二，栅格 GIS 的叠置分析中不存在"空间操作与分析"两层次差别，换言之，栅格型叠置分析的操作就是分析，因为栅格型叠置操作就是基于各叠加图层像元的属性数据，按某种关系，确定叠置结果层中相应像元的属性值：

$$E_{ij} = f(P_{ij}^1, P_{ij}^2, \cdots, P_{ij}^n) \qquad (5-2)$$

式中，E_{ij} 为叠置结果层第 i 行第 j 列像元的属性值；$P_{ij}^1, P_{ij}^2, \cdots, P_{ij}^n$ 分别为参与叠置的第1，第2，…，第 n 层的第 i 行第 j 列像元的属性值。

这样，叠置操作的本身实际上已经完成了式(5－1)所做的分析评价工作。例如，若参与叠置的图层为土地利用、土壤和坡度因子层，所采用的分析关系 f 与式(5－1)相同，则叠置操作式(5－2)的结果，事实上业已给出该土地评价的结果。因此，栅格型叠置操作和分析是一回事。

无论是采用栅格型叠置还是矢量多边形叠置进行分析评价，皆可以采取各种各样的叠置分析运算关系(即上两式中的 f)。运算关系可能来自理论模型，也可能是经验公式。从数学形式上讲，叠置运算关系可以是一般的数学运算，如加、减、乘、除、乘方、指数、对数、三角变换等，也可以是逻辑运算，如与、或、非、异或等，或其他运算。

实践中，几种运算常联合使用。例如，某自然保护区欲建设一商业中心，在为之进行的环境评价中，需要计算该项建设引起的地面径流量增加(关系到水土保持)；而计算径流量与土壤类型有关。为此，计算径流量的程序中，既利用水文学公式，又对土壤类型实施逻辑运算，即如果土壤类型＝…时，那么…

在栅格型叠置中，二值逻辑叠加也是一种常见形式。二值逻辑叠加常作为栅格结构的数据库查询工具。例如，比较复杂的查询涉及多种复合条件，如查询所有的面积大于10 km²，且邻近工业区的全部湿地。这种数据库查询通常分为两步，首先为每个条件创建一个新图层，通常是二值图层，1代表符合条件，0表示不符合条件；第二步进行二值逻辑叠加操作得到想查询的结果。

5.2.3 缓冲区分析

1. 缓冲区分析的概念

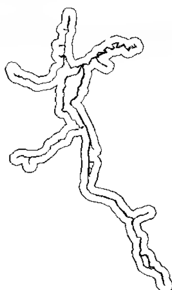
缓冲区概念已见于§5.1.3.1，指与点、线或多边形边界的距离为给定值的区域。在§5.1.3.1中说过，该处仅简述缓冲区概念，且所讲的缓冲区查询仅以查询为主要目的，所生成的缓冲区是暂时的，不改变原文件，且不一定要求是拓扑数据。

本节所讲的缓冲区分析则不同。首先，它对一组或一类地物按缓冲的距离条件，建立缓冲区多边形图，可形成缓冲区图文件；第二步，将这一个图层与需要进行缓冲区分析的图层进行叠置分析，得到所需要的、并可保存的结果。缓冲区分析也称为缓冲区操作；但有时也将上述第一步称为缓冲区操作，第二步称为缓冲区分析。

缓冲区有点缓冲区、线缓冲区和面缓冲区之分。一个点状地物的缓冲区为圆形。线缓冲区形状如图5－4所示，该图是一条河流的缓冲区，由于该河流部

序号	土地利用类型	面积 亩
1	灌溉水田	4 856 6
2	旱地	2 174 6
3	果园	232 4
4	桑园	151 8
5	茶园	885 1
6	有林地	10 951 6
7	居民点用地	2 137 6
8	交通用地	203 2
9	水域	555 0
10	未利用地	612 9

(a) 夏履江部分河段 200 m 范围内各类土地利用面积统计



(b) 夏履江部分河段 200 m 缓冲区

图5 4 一个缓冲区分析的实例

分双线表示,该缓冲区实际上有一部分是面缓冲区。注意,面缓冲区的缓冲距离,是相对于面状地物的边界而言的。因此,面缓冲区有外缓冲区和内缓冲区之分:前者是面状地物外围形成的缓冲区,后者是在面状地物边界内侧形成的缓冲区。当然,也可以同时形成内、外缓冲区。

无论是点缓冲区、线缓冲区还是面缓冲区,皆可以进行缓冲区分析,求取缓冲区内的点状、线状和面状地物的信息,因为,点、线和多边形的缓冲区都是多边形,都可以通过"点与多边形叠置"、"线与多边形的叠置"和"多边形叠置"的操作,分别与点状、线状和面状地物的图层进行叠置分析。

缓冲区分析是 GIS 中典型的邻域分析,它描述地理事物的一种影响范围或服务范围,具有重要的应用意义。例如,欲求取公共设施,如商场、邮局、银行、医院、车站和学校等的服务半径,基本建设,如道路和大型水库建设等所涉及的搬迁范围,铁路、公路以及航运河道对其所穿过区域经济发展的重要性,以及点状、线状和面状污染源的影响范围等,都可以进行缓冲区分析。

2. 缓冲区的生成

从原理上说,缓冲区的建立在技术上比较简单。点缓冲区以点状地物为圆心,以缓冲区距离为半径绘圆即可;线状地物和面状地物的缓冲区的建立,是以

线状地物或面状地物的边界为参考线作其平行线，同时在端点和拐点等处绘制圆弧，与平行线相接即可。

但是，在实际的技术实现中情况比较复杂，因为缓冲区往往会重叠，边界可能相交。解决此问题可采取两种途径，一是在缓冲区生成过程中解决，即在作参考线的平行线时，考虑各种情况，确定相互间的交点，切断并清除重叠区内的弧段。二是在缓冲区生成后解决，即判断各缓冲区的重叠部分，对重叠部分进行多边形合并，清除缓冲区外界内的相交弧段。此外，当生成的缓冲区较多、较密或弯曲婉转时，有些缓冲区与非缓冲区的多边形难以区分，在处理时需要给定标志予以区别。

在建立缓冲区时，缓冲距离可以根据地理对象的特征合理给定，必要时可以因地而异。例如，沿河流给出的环境敏感区的宽度应可根据不同河段特征而调整；也可以对同类地物，一部分建立、一部分不建立缓冲区。这时可以扩展属性表，增设字段并给不同地物指定不同的缓冲区距离。缓冲区的形态也可以灵活变化，如点对象可有三角形、矩形等形状的缓冲区；线对象可有双侧不对称或单侧的缓冲区等。事实上，虽然通常意义上的缓冲区概念是等距离扩张的，但也不必拘泥，因为建立缓冲区的宗旨是进行邻域分析，等距离扩张的邻域不一定是符合实际情况的邻域分析范围。

3. 缓冲区分析实例

图5-4给出一个缓冲区分析的实例，这是浙江省绍兴县夏履江小流域的一部分。在流域规划中，欲对夏履江部分河段进行改造施工，需要了解可能波及的200 m范围内，有多少农用地、果园、林地河居民用地等，以辅助规划决策。为此需要进行缓冲区分析。首先，对夏履江有关河段建立缓冲区；然后，将如图b所示的缓冲区与该地区土地利用现状图层进行多边形叠加操作并进行查询，得图a所示的表格数据结果。

5.2.4 网络分析

1. 网络与网络分析的概念

网络(network，或circuits network)是现实世界中，由链和结点组成的、带有环路(loop或circuit)，并伴随着一系列支配网络中流动(flow)之约束条件的线网图形。"网络分析"是通过模拟、分析网络的状态以及资源在网络上的流动和分配等，研究网络结构、流动效率及网络资源等的优化问题的领域。现实世界中，大量物质的运行或传送是在网络系统中进行的。有形的网络系统如：各种水、陆、空的交通网络及其中的旅游、配送、急救等流动；具有环路的河网和灌溉网；各种城市管网，如排水(雨水、污水)、给水(自来水)和煤气的管网，以及电力、电信和有线电视的线路网等。无形的网络系统如微波或无线通讯网等。现

实世界如此多的网络系统及资源在其中的分配、流动的不同特性，以及各种支配、约束条件等，使网络分析既具有丰富的内容，又有重要的研究和应用意义。

进行网络分析研究的数学分支是图论和运筹学，研究内容包括最佳路径、资源分配、结点或弧段的游历（旅行推销员问题、中国邮递员问题）以及最小连通树、最大、最小流等问题。GIS 技术兴起后，由于网络分析的重要意义，GIS 凭借自己的地图分析能力和通过地图表达和分析地理世界中网络系统及其流动的优势，又从图论和运筹学吸取营养，依据网络拓扑关系，包括链和结点之间的连接、连通关系，通过考察网络元素的空间及属性数据，对网络的性能特征进行多方面的分析计算，从而发展起自己的网络分析技术。 部分具有普遍应用意义的网络分析手段，被进一步总结、提炼出来，形成了 GIS 软件中的网络分析功能或模块。目前，GIS 网络分析在电子导航、交通旅游、各种城市管网和配送、急救等领域中发挥重要的作用。

网络数据结构的最基本组成部分，一是结点（node），它体现网络中的连接、连通关系；二是 link，这里译为"链"，它是网络中每两个结点之间的线形单元。在网络中，中心（center）是接受或分配资源的、具有战略意义的结点，如水库、商业中心、电站等；站点（stop）是路径选择中资源增减的结点，如库房、车站等。网络必须具有环路，否则只是单纯的"树"结构（如河流及其支流）。网络还带有约束条件。例如，如自来水的水压在距处理中心一定距离处将为零，一个学校的容纳或招收学生的能力总有限度等，此即服务范围限制；网络中还可能有"障碍"（barrier），即禁止网络中链上流动的点，以及"拐点"（turn），即限制方向流动，例如不允许左拐等。GIS 网络分析按照给定的应用需求，在一定约束条件下，基于上述结构和网络关系的分析计算，寻求最优解决办法。

GIS 网络分析主要有最佳路径、资源分配、服务中心选址和地址匹配等问题。地址匹配已在空间查询部分简单介绍，下面仅讲解前三个问题。

2. 最佳路径

最佳路径要解决的问题是：What is the best path? 解决问题一般有下述三个步骤。

第一步，根据应用目的确定"最佳"的含义。按"地理空间"的扩展概念（参照 §5.3.2.1），地理空间的距离有多种含义。从社会应用目的角度，空间距离可分为很多类别，例如：

- 欧氏（欧几里得）距离。即普通意义上的距离。
- 时间距离（Time distance）。体现从空间一点至另一点旅行或流动所花费的时间。
- 经济距离或成本距离（Economic distance or Cost distance）。体现从空间一点至另一点旅行或流动所花费的经济代价或成本。

- 风险距离(Risk distance)，体现空间一点至另一点旅行或流动的风险大小。
- 认知距离(Cognitive distance)。体现人类认知中的空间两点间的距离。
- 社会距离(social distance)。体现空间两点间的社会亲和程度。
- 生态距离(Ecological distance) 体现空间两点间的生态关系。

上述空间距离的概念中,网络分析中常用的距离概念是欧氏距离、时间距离、经济距离和风险距离。其中,又以时间距离和经济距离最为常用。很多网络分析以流动最快的路径为最佳路径;也有很多的网络分析在流动时间差异不大的条件下,以经济距离最小的路径为最佳路径。有的网络分析则将风险因素放在第一位。

另一方面,最佳路径概念还涉及分析计算。从计算角度,空间距离可分为:

- 直线距离。即按两点间直线计算的距离。
- 旅行距离。为沿地理空间中限定线路计算的距离(此概念也适用于非网络的线路),因为大量旅行和流动是沿限定线路,如道路、河流或管网等进行的。在计算上,旅行距离等于沿旅行路线的若干小段距离之总和。在城市道路网行进的旅行距离,有时称为街区距离。
- 统计距离。即统计学意义上的距离。统计距离常用于聚类分析。在空间变量多维空间中,统计距离体现具体样点归属某具体类别的可能性大小。由于不同类别样点分布差异,基于统计距离的分类计算,常常比基于欧氏距离更为有效、合理。这一点我们将在 § 5.3.1.4 中具体讲解。

上述距离中,网络分析比较常用的是旅行距离,因为网络中两点间的流动一般沿线路进行。只有航空网络例外,它一般计算直线距离,因为直达的航班沿地表直线(大圆弧)飞行。

综上所述,网络分析中的最佳路径,一般指上述某种特定含义下,距离为最短的路径。例如,若某网络分析采用时间距离概念,并按旅行距离计算,那么,最佳路径就是沿某给定网线流动时间花费最少的路径。但也有少数网络不一定采用最短距离或时间。例如电网,由于电的传输速度极高,它考虑的角度有些不同。尽管如此,绝大多数网络分析采用最小距离的最佳路径概念。

上述空间距离的内容也适用于后面的资源分配和中心选址分析。

最佳路径分析的第一步,是在给定某种距离概念的基础上,计算确定从任意给定起点和终点间的最佳路径。不难理解,网络中给定起点和终点间可能有很多条通路。必须计算每一条可能的有效通路的距离,该距离通常是通路中每条链的距离之和。例如,某起点和终点间的某条通路的时间距离,是该路径每条链所需时间之总和。每条有效通路的距离计算出来后,进行比较,从中找出其中的距离最短者,即为最佳路径。

第三步，将最佳路径贮存在属性表中。其目的有二，一是能随时在地图上可视化显示；二是可用于进一步的网络分析，如资源分配和中心选址分析中。

在最佳路径分析的上述三步骤中，从技术上讲，较难处理之处是如何提高从起点至终点各种可能通路的搜索计算效率。在复杂网络中，从起点至终点的可能路径太多，从相反方向出发绕一大圈达到终点也是一条路径。为此需要研究一些较好的算法。例如，E. W. Dijkstra 于 1959 年提出的一种基于有向图的两点间最短距离的算法，为 GIS 广泛采用，有兴趣的读者可参见龚健雅或邬伦的 GIS 教材。

为了计算非欧氏距离，需要给定每条链以相关属性（或权值）。所以，非欧氏距离也可以说是一种加权距离。例如，在美国的交通网络中，无论数据库内还是道路两旁的醒目标志牌上，都有每条道路的限定速度；且车辆普遍遵守规则大体按限速行进（过慢也违法）。这样，限定速度就可以用来作为计算时间距离的参数（或权值），因为某段路（即链）的"限速"乘以该段路的旅行距离，即为该段路的时间距离。

此外，在实践中，除了理想情况的最佳路径外，有时还需选择若干近似的最佳路径以备用；有时实际网络中权值需随某种关系而变化，还可能会临时出现一些障碍点，需要动态计算最佳路径等。

3. 资源分配

资源分配要解决的问题是 What center serves a particular link? 进行资源分配分析的网络中，一般已有若干服务中心，要求优化划定每个中心的服务范围，即解决每个特定的链由哪个中心来服务最为合适。这显然是一个优化配置网络资源的问题。例如，城市中有若干投票站，为了最大限度方便公民投票，市政当局可通过资源分配分析，确定每条街道（道路网中的"链"）划归哪个投票站，并据此给公民发函，请公民到那个投票站。周到的政府还在通知函上绘有从该街道到投票站的地图。

资源分配分析的步骤是：

第一，按上述计算最佳路径的方法，计算出网络中某一条参与运行的链到每个中心的最佳路径，并进行比较，找出最佳路径的距离最短的那个中心，便将该链划归该中心。例如，采用时间距离概念，某街道到投票站 A、B、C 和 D 的最佳路径的时间距离，分别为 15、17、25 和 34 min，则将该街道划归投票站 A（当然，用真正民众公仆的话，应当说是指定投票站 A 服务于该街道）。

第二，对网络中每条链重复同样工作，直到将所有的链都归于某个中心为止。当然，划分归属区必须符合约束条件，如水处理中心在一定范围处水压为零，归属区不能超过这个界限。为此，供水能力强的中心可能要扩大服务范围；或者改建或增补该区域的供水设施。

第三,将划定的服务中心的标识贮存到链的属性表中,以便能随时检索和显示分配给每个中心的所有链的范围、区域),同时也可为进一步分析所利用。

资源分配分析具有重要的应用意义。城市和区域中有大量的服务中心,除很多公益型的中心如邮局、学校、医院和急救站等外,商业服务中心更多,都可以通过资源分配分析来提高效率。

4. 服务中心选址

服务中心选址是也是资源分配分析的一部分,是上一问题的延伸,它要解决的问题是:What would locate a service center? 由于任何建设都不可能随心所愿地选择地址,在服务中心选址中,一般给出几个可供选择的地址,进行比较分析。在服务中心选址中,约束条件(如欲建学校可容纳学生的最大数量等)通常更为重要。服务中心选址要求在给定约束条件下,对每一服务中心的选址方案进行资源分配计算,求取平均而言或总体而言的最佳方案。

服务中心选址的步骤是:

第一,对各种可能的选址方案进行上述资源分配分析。在这种分析中,必须将规划建设的服务中心与现有的中心合在一起通盘分析。有时,分析可能表明兴建一个中心达不到较好的效果,需要考虑在相似经费框架下分两处兴建。这样,规划中心与可选地址都不止一个,它们之间的配置情况,即选址方案将会有较多种。不论怎样,每一种选址划方案必须加上已有的中心,通盘分析。例如,某城市已有4所小学,欲向西北扩展,在3个可选地址上兴建2所条件不同的小学,则会有6种排列,即6种选址方案。不论哪种方案,都要将规划的2所小学与原有的5所小学合起来通盘分析,即:如果有了如此7所小学,资源分配情况会怎样。对6种选址方案都进行资源分配分析,划分服务区,进行不同方案的显示。

第二,对每种选址方案的资源分配或服务区划分结果,计算该方案中所有参与运行的链的网络运行距离的总和或平均值。例如,上例中,对6种选址方案中的每一种的(7小学)服务区分划,计算全市所有街道的适龄学生(需要有详细人口数据)赴学校的时间距离之总和或平均值。

第三,比较每种方案,选择上述总和或平均值为最小的方案。例如,假定6种选址方案中,别的方案全市所有学生上学平均时间皆在6 min以上,只有一种方案(例如较大容量小学建在西南,较小容量小学建在北边的选址)实施后,全市所有学生上学平均时间为5.5 min,同时该方案又能满足建校约束条件,那么,该选址方案将作为最优方案采用。

当然,实际应用非常复杂,最后的选择还可能要考虑很多其他因素。例如,小学校的选址问题,还应考虑学生穿越马路的安全性;对于商场的选址,要考虑交通状况,分划区人们的经济能力、消费水平、文化素质等因素;电信移动通讯

的基站选址不但要基于其专业模型,还要考虑地形起伏 建筑物的遮挡等因素;在不少管网系统中,最佳路径常需要考虑网络传输的单双向、限制转向等问题;商贸中心还需要注意发货中心和收货中心的区分;最佳路径常需要综合考虑多种因素,例如选择旅游路线需要综合考虑时间、经济、兴趣和安全等多种因子,单从一种角度,如时间距离或经济距离来进行网络分析,虽然相对简便,实际上却是不够全面的。

上面的讨论仅限于可以看做是 GIS 软件中程式化的、且较为典型的网络分析功能。事实上,网络分析所涵盖的内容,包括运筹学研究的内容还很多。

§5.3 其他空间分析

本节介绍尚未谈到的各种空间分析。第一分节的内容大体相当于一般 GIS 书上的"空间统计分析"。其中某些写法,例如空间统计分析与一般的多变量统计分析的区别,以及空间统计分析与栅格数据处理方法内在联系,带有作者的观点。第二分节介绍与分析地图学有关的空间分析概念。第三分节介绍具有"提高"性质的空间分析。鉴于本书的定位,第二、第三分节皆只作较初步的讲解。

5.3.1 空间统计分析及其与栅格数据处理方法的联系

1. 空间统计分析的内容和意义

空间统计分析主要用于地物分类和综合评价。分类是科学研究的一种基本方法。当遥感数据等地理空间数据大量采集和涌现时,首先需要从特定学科(如植物、水文或土地资源等)角度进行科学分类;分类的方法和质量,对后续空间数据处理有明显影响。进而,人们需要基于多学科或多领域数据,对各种资源和环境进行综合评价,以求较佳的资源配置和环境质量。在上述这些工作中,常常基于数理统计方法进行空间统计分析。

众所周知,现代数理统计已在大量的学科和应用领域发挥十分重要的作用,很多数理统计手段业已成为通用软件(如微软的 Excel 等)和不少领域专业软件的现成功能。GIS 领域自不例外。由于地理空间变量的多样性、复杂性和海量数据的特点,GIS 在统计学理论和方法的应用上,不仅比很多其他领域更加深入,而且形成了自身可贵的特点。本小节所谈的空间统计分析,就是指 GIS 系统中、基于计算机化的数理统计手段所进行的多个空间变量的综合分析。

空间统计分析的基本方法与一般多变量统计分析类似。首先,要有多个变量,每个变量要有一串样本数据。为了进行多变量统计分析,通常将多变量的多样本组成一个数据矩阵。如表 5-2 所示,假定有 n 个变量、m 个样点数据,

则可组成一个 n 行、m 列的数据矩阵(该表是为帮助理解而特别给出的):

表 5-2 空间变量所组成的矩阵

具有各自地理位置的		样点 1	样点 2	样点 j	样点 m
空间变量	X_1	x_{11}	x_{12}	x_{1j}	x_{1m}
	X_2	x_{21}	x_{22}	x_{2j}	x_{2m}
	X_i	x_{i1}	x_{i2}	x_{ij}	x_{im}
	⋮				
	X_n	x_{n1}	x_{n2}	x_{nj}	x_{nm}

表中可见,数据矩阵的每一行代表一个变量,每个变量都有 m 个样点数据;另一方面,每一列代表 1 个样点,每个样点都各有 n 个变量的数据。基于数据矩阵,人们常采用多维向量空间概念:n 个空间变量形成 n 维空间,每一个样点的多维数据(表 5-2 数据矩阵中的一列),组成向量空间中的一个向量。基于多维向量空间概念和线性代数方法,众多有关学科都演绎出有自身特色的、内容丰富的结果。

那么,空间统计分析与一般多变量统计分析有什么区别呢?根本区别在于每个样点都有各自地理位置特征,这样的变量才是空间变量。换言之,空间统计分析的每个样点体现着一个地点或地理事物。例如在城镇土地估价中,要求知道城镇每一个地块的一系列空间变量,例如,有形的区位因子:环境质量、人口密度、道路网密度和生活设施完备度等,以及无形的区位因子:商服繁华度、道路通达度、公交便捷度和社会服务设施完备度等;每一个地块相当于表 5-2 的"样点",但每个样点都有地理位置特征。

空间统计分析所要进行的地物分类和综合评价,都是在 n 个空间变量 X_1,X_2,X_3,…,X_n 的综合分析基础上,求取一个或多个新的空间变量(即表 5-2 增添新的行)。在地物分类的情况下,新的空间变量就是分类结果;在综合评价的工作中,新求得的空间变量就是评价结果。这些结果通常用数值表示。例如,某样点分类结果值为 5,表示该样点属于第 5 类。当然,在求取分类和评价结果的过程中,常常需要求取很多的中间变量,或辅助变量。不论得到哪种新变量,组成它的样点及其地理位置保持不变。

由此可知空间统计分析与一般多变量统计分析的另一个重要特点,这就是:基于每个样点的地理位置特征(地理坐标或体现地理分布的其他坐标等),空间统计分析的结果可以进行地理可视化表达。例如,图 4-12 就是一个土地估价结果的三维可视化表达。当然,一般还是采用二维方式,即在平面地图上

利用不同色调，表达每一个地块的分类或评价分析结果。可视化对进一步应用或研究具有重要价值。

2. 空间统计分析与栅格数据处理方法的联系

从表5-2可见，空间统计分析可以基于一个数据矩阵和多维空间等手段，采用矩阵分析方法；而我们记得，栅格数据处理方法也主要是矩阵方法。因此，空间统计分析与GIS栅格数据处理之间有不少互通或可借鉴之处。我们说过，GIS中最重要的栅格数据处理是遥感图像处理，最重要的栅格GIS软件就是遥感图像处理软件。因此，这里有必要说明空间统计分析与遥感图像处理的联系。

遥感图像处理最重要的内容是基于多波段遥感数据的分析和信息提取。在这种分析中，每一波段的图像数据相当于一个空间变量；每一个像元就是一个样本，皆带有其地理位置数据（地理坐标或行列序数）。每一波段全部像元之值组成该波段（空间变量）的样本数据。遥感图像处理就是通过各种变换，在每个像元上产生新的数值，即生成新的图像，包括分类图像或综合评价图像。因此，遥感图像处理软件中的不少功能其实就是空间统计分析功能。

应当特别指出，自20世纪70年代陆地卫星数据出现以来，世界各国在遥感应用上投入大量经费，众多的遥感应用领域的大量专家学者进行了数十年遥感图像分析处理研究，动用了一切可能的计算机统计分析手段，在应用现代数理统计手段从海量数据中自动提取信息的研究上，其规模之大，其钻研之深，为其他很多计算机统计分析的应用领域所不及。特别在自动分类研究方面，遥感图像处理无疑走在科学界前列。空间统计分析可以从遥感图像处理技术及软件借鉴很多的东西。

下面将分别介绍地物分类和综合评价的空间统计分析方法，以及为它们服务的空间变量筛选问题。其中某些思路或写法是从遥感图像处理中借鉴来的。

3. 空间变量筛选

空间变量筛选又称为变量分级（ordination），是空间统计分析的一项重要的预处理工作，无论是地物分类和综合评价，都常常需要进行这样的预处理。

空间变量筛选的必要性，主要由于很多空间变量之间有信息冗余。为了理解这一概念，我们首先来讨论一下数据中所含信息与方差、协方差（相关系数）的关系。一般而言，一个空间变量所含信息量的多少可用其方差表示

$$V_i = \sum_{j=1}^{m} (x_{ij} - A_i)^2/(m-1) \qquad (5-1)$$

式中，V_i 和 A_i 分别为第 i 个空间变量的方差和平均值；m 为样本总数。

从式（5-1）可见，方差是空间变量的样本值 X_{ij} 距离其平均值的偏离之总和，它反映其样本值的“波动”程度。一般而言，波动幅度愈大，所含信息量愈

多;反之则愈少。例如,若某波段遥感图像的像元值都在平均值附近(波动很小),不难想象,该图像将是一幅没什么明暗变化的"灰蒙蒙"图像,所含信息很少。

另一方面,两个空间变量,例如,第 i 个空间变量与第 k 个空间变量(平均值分别为 A_i 和 A_k)之间的信息冗余程度,可用二者之协方差或相关系数表示。协方差公式为:

$$V_{ik} = \sum_{j=1}^{m} (X_{ij} - A_i)(X_{kj} - A_k)/(m-1) \qquad (5-2)$$

协方差和相关系数是性质相同的两个因子,差别只在于相关系数要在协方差公式(5-2)的分母中加一项归一化因子,以便能分别以1和0来表示全相关和不相关。在空间统计分析中,当不讲究归一化时,为提高计算效率,通常采用协方差。

协方差或相关系数为什么能体现信息冗余呢?从式(5-2)可见,当两个变量对其平均值的偏差"同步"变化,即一个为正(或负)偏差时,另一个也为正(或负)偏差时,协方差便较大。这种"你大我也大,你小我也小"的两个变量,正是我们说的"非常相关"的同类因子,它们所反映的,很可能是同性质的空间变化信息。如果在空间统计分析中不加分析地同时采用多个同类因子,那就夸大了同类因素的贡献,这就是"信息冗余"。

例如,遥感图像中的两个可见光波段的明暗变化通常比较类似,一个图像上较亮(暗)之处另一图像上也较亮(暗),二者之间就有较大的信息冗余。又如上面提到的土地估价因子,如商服繁华度、道路通达度、公交便捷度和社会服务设施完备度,就明显有大量冗余,因为在发达(或偏远)的地方,这些因子的数值一般都偏高(或偏低)。因此,同时采用这些因子并给予不小的权重,是不合适的。

空间变量筛选或分级的任务,就是尽可能减少信息冗余。筛选的方法较多。简单的办法是直接利用协方差或相关系数,结合经验,剔除或合并一部分因子。较科学的筛选方法之一,是主成分分析。

主成分分析是一种数学方法,它基于变量间的协方差矩阵,构建(线性)特征方程;以解出的特征根进一步推求特征向量;以特征向量组成的矩阵作为转换矩阵,将原有的 n 个空间变量,转换为另外 n 个新空间变量。按线性代数严格证明,这样的转换结果将有下述特点:

第一,所有新变量之间的协方差全部为零,即没有冗余;

第二,新空间变量的方差之总和,与原空间变量的方差之总和相等,即"信息"总量不变;且主要方差集中在少数几个新变量上。

新变量按方差从大到小,被分别称为第一,第二,…,第 n 主成分。这样,主

成分分析就将众多要素的信息压缩、表达为若干具有代表性的合成变量，以致进一步的空间统计分析可只取方差最大的若干个新变量来进行。这就克服了变量选择时的冗余。

下面虽然未写出主成分变换的数学过程，但这里以二维空间为例，给出可视化解释。如图5-5所示，假定两个空间变量 X_1 和 X_2 的样点在二维空间呈椭圆形分布，且愈近椭圆中心，样点愈密（这是典型的正态分布，银河系很多球状星团亦如此）。X_1 和 X_2 之间有一定协方差，因为总的来说 X_1 和 X_2 有共同增减的趋势。但是，若将坐标轴旋转一个如图所示的角度，即通过线性变换（坐标平移和旋转即线性变换）转换成两个新变量 X'_1 和 X'_2，则同样是这些样点，在新坐标系下两个新变量之间的协方差将为零；而方差将集中到位于椭圆长轴方向的变量 X'_1 上。

图5-5 主成分分析示意

当然，主成分分析并不是完美无缺的，数学上的理想并不能保证理想的实践效果。地理事物千差万别，统计学上的理想结果远不能代替具体地理特征的分析。

4. 地物分类

常用的地物分类方法主要有聚类分析和判别分析两类。其中，聚类分析更为重要，遥感图像处理的分类技术为传统聚类分析增添了不少的手段和内容。例如，早几年某些英文 GIS 教材中的聚类分析（clustering），仅相当于遥感图像分类技术中非监督分类方法。

聚类分析将多个空间变量看做多维向量空间，每个样点的多变量数据（参见表5-2及其说明）决定了样点在多维空间中的位置。例如在图5-5中，每个样点的 X_1 和 X_2 两个坐标值，决定该样点在 X_1 和 X_2 两变量的向量空间中的位置。将所有的样点都点绘在二维向量空间中，就形成图5-5所示的样点空间分布。

多维空间虽然不能像二维空间那么好想象，但道理是一样的。聚类分析的基本思想是：将多维向量空间中位置接近的样点划归一类；其准则是：使类间差异尽可能大，而类内差异尽可能小。

判断位置"远近"的距离，可以有不同的含义。一种距离是欧氏（欧几里得）距离，包括直线和非直线的欧氏距离，每段距离按两点坐标分量之差的平方之和计算（因聚类分析只比较远近，无须计算平方根）。聚类分析中常采用另

种距离,是统计学意义上的距离,我们称为统计距离。统计距离体现具体样点归属某具体类别的可能性大小,采用统计距离的聚类分析一般比采用欧氏距离科学。这是因为,不同类别样点在向量空间中的分布范围或弥散度通常差异较大,按欧氏距离将样点划归离它最近的类别,并不一定合理。

让我们用图5-6来加以解释。如图所示,仍假定在二维空间中分类,图中 A 和 B 两个椭圆分别表示两个类别样点的大体分布区。例如,A 为某种山丘地样点的大体分布区;B 为湖泊样点的大体分布区。现在图中有一个样点 X,按欧氏距离,它应当划归 B 类,因为 X 至 B 类中心近得多。然而,按统计距离判断就不同了,因为 A 类本身分布的弥散度较大,B 类样点集中在小范围内。这样,X 属于 A 类延伸过来的可能性反而较大;属于 B 类的可能性反而小。统计距离就能体现归属的可能性。但统计距离一般要求样点呈正态分布,这是它的一个缺点。

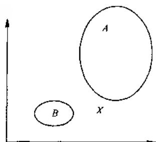


图5-6 统计距离的作用

统计距离有多种,这里就不一一解释了。

另一种地物分类方法是判别分析法。它与上述聚类分析不同:聚类分析的结果可以是地物的类型,如植被、水域、城镇等,也有时是地物的等级序列,如土地分等定级、水土流失强度分级等;但判别分析仅用于等级序列分类,它建立在下文所讲的综合分析评价的基础上,最后采取措施判别等级。因此,我们将它归于下一小节内容。

无论是地物分类还是下文的综合分析评价,从人类认识的角度来看,都有精确的和模糊的两种类型。由于绝大多数地理现象难以用精确的定量关系划分和表示,模糊分类和评价模型常更为实用,结果也往往更接近实际,因而在地理信息系统中较多地得以发展和应用。在模糊分类和评价模型中,首先确定各因素各类别对目标的支配隶属度,作为判别距离的度量;再结合要素的权重指数,采用适当的模糊算法,计算各地理实体的归属或等级类别,作为进一步分析评价的基础。该方法通过隶属度表达人们对目标与因素之间关系的模糊性认识,用适当的算法将这种认识量化,并反映到结果的分类中;算法依据不同因子的变化情况确定,常采用分段线性函数或其他高次函数。

5. 综合分析评价

综合评价模型是区划和规划的基础。综合评价一般经过四个过程。

第一,评价因子的选择与筛选。分析评价通常会涉及大量相关因子,应根据理论和经验认真细致地分析、选择,并进行上面讲过的因子筛选工作。

第二,多因子重要性指标(权重)的确定。因子权重的确定是建立评价模型的重要步骤,权重正确与否极大地影响评价模型的正确性。

第三,因子内各类别对评价目标的隶属度确定。这是采用模糊模型的需要

第四,选用某种方法进行多因子综合

鉴于因子权重确定的重要性,为了避免确定因子权重对主观判断的较多依赖,一种综合众人意见,科学地确定各影响因子权重的简单而有效的数学手段被发展出来,这就是层次分析法。

层次分析(analytic hierarchy process,简称 AHP)是系统分析的手段之一。复杂系统通常涉及大量相互关联、相互制约的因素,决定诸因素对目标重要性的序列,对建立模型十分重要。层次分析就是要解决这个问题,它把人的思维过程层次化、数量化,采用定性和定量分析相结合的方法,把相互关联的要素按隶属关系分为若干层次,请有经验的专家对各层次各因素的相对重要性给出定量指标;利用数学方法综合专家意见给出各层次各要素的相对重要性权值,作为综合分析的基础。

例如,在旅游问题中,假设考虑 5 个因素:费用 y_1、景色 y_2、居住条件 y_3、饮食条件 y_4 和旅途条件 y_5。为了比较 5 个因素对旅游目标的影响,确定它们在分析评价中的比重,可每次取两个因素,请有经验的专家进行比较。通过这种成对比较法,得到一个正互反阵:

	y_1	y_2	y_3	y_4	y_5
y_1	1	2	7	5	5
y_2	1/2	1	4	3	3
y_3	1/7	1/4	1	1/2	1/3
y_4	1/5	1/3	2	1	5
y_5	1/5	1/3	3	1	1

矩阵中各元素之值体现它所在行、列的两个因素对分析评价的重要性之比。例如,1 行 2 列元素等于 2,表示费用因子 y_1 与景色因素 y_2 对选择旅游目标的重要性之比为 2 比 1;1 行 3 列元素等于 7,表示费用 y_1 与居住条件 y_3 的重要性之比为 7 比 1;3 行 2 列为 1/4,则表示居住条件 y_3 与景色因子 y_2 的重要性之比为 1 比 4 等

有了上述矩阵,就可以用来决定 5 因子旅游评价中 5 个因子的权重。

在综合分析评价的基础上,可以进行上面提到的判别分析法。判别分析的

前两个步骤与综合评价相同,即选择若干评价因子并进行线性组合,确定各要素的权重和隶属度。再接下,判别分析要求构成一个线性判别函数,对样本进行判别。然后,将判别分析的结果序列按等间距或不等间距的标准划分,计算出判别临界值;最后,判别每个样点最可能的评价等级。可见,判别分析法是一种带有分类和综合评价双重特点的空间统计方法。

5.3.2 分析地图学与空间分析

地图分析具有悠久的历史。直到20世纪50年代,地理分析始终从欧氏空间的角度,绝对化地对待"地理空间和距离"的概念。从60年代起,随着计算机制图和GIS的出现,人类的地图分析能力极大加强,导致"分析地图学"诞生(Tobler,1961)。分析地图学可以简单地定义为数学的、分析的地图学,它现在已发展成为一个与GIS有较大交叠的学科领域。本节介绍与分析地图学有关的空间分析概念,包括基于矢量地图的空间统计分析,以及基于地理空间和距离拓展概念所进行的空间分析;最后介绍现代分析地图学。本节的介绍均为简单介绍,且仅限于基本概念,主要目的是扩大视野。

1. 基于矢量地图的空间统计分析

基于矢量地图的空间统计分析是传统地图分析的一种延展。由于地图进入计算机,人们得以实现过去在手工方式下难以进行的、地图图元空间分布特点的分析,很多数理统计手段被引入这个领域。1981年,David Unwin在总结不少人工作的基础上,出版了这方面的代表著作Introductory Spatial Analysis。下面我们就以该著作为素材简介这一类空间分析。

基于矢量地图的空间统计分析在论述地图及其建模,以及地图分类学的基础上,从传统的几何或形态学角度,以数理统计为主要分析手段,对地图上的点、线、面、体的地图建模和空间分布特点进行分析。地图上的"点"的分析内容主要包括点符号绘制的理论实践和点的分布特征(pattern)。后者又从三种角度讨论:一种角度是基于点分布的密度;第二种基于距离(欧氏距离或时间、经济距离等);第三种基于距离和方向。无论哪一种,一般都设法定量化为矩阵表达,以便进行统计分析。

以基于密度的点分布为例。它不是单纯的密度计算,其主要目的是分析地图上点的分布密度是否具有统计学意义,分析其空间依赖性。一种典型方法是方格分析(quadrant analysis),即用一个方格网覆盖地图,以落入每个方格的点的数目组成矩阵,作为统计分析的数据;然后用各种可能的统计检验方法判断该分布的置信度等。

地图上的"线"的分析内容,除线符号绘制的理论实践外,线被分成三个层次:路径(path)、树结构(tree structure)和网络(network)。树结构和网络都是由

链(相当于单个 path)和结点组成的线结构,但网络中有环路(已见§5.2.3),而树结构没有。典型的树结构是流域水系(但一些平原河网和灌溉网中有环路)。

路径又分为单个路径和路径群。对单个路径,可以进行长度、走向、蜿蜒度、分形几何和谱(傅立叶)等的分析;可以建立地图中单个路径方向变化的矩阵等。对路径群,即密集的路径,可分析平均长度、标准差、路径密度和频率(单位面积内的 path 数目)等。

对树结构,可对组成它的链(如河流)分级,以最上游为1级,1级链汇入2级,2级链汇入3级等;可以发现每一级链的数目分布具有一定的规律。还可以分析树结构中的链的长度、结点处链的角度,以及结点的空间分布等。

对网络结构的地图学分析不大同于§5.2.3的网络分析。一种常用的方法称为 graph 方法,该理论将 graph 看做是网络或树的结点及结点间关系之集合。这些关系包括:两点间对等且可逆,如两城市有路相通;可逆但不对等,如容量不等的两台互连的交换机;两点间不可逆,如单行道连接的两点等。graph 方法以不同编码表示不同关系,组成结点的关系矩阵(其形态与上述§5.3.1.4中矩阵类似),再进行统计分析。

对地图上的区域(area、二维)和表面(surface,三维)图元也有很多统计分析方法,这里就不一一介绍了。上面叙述了一些关于地图上点和线的分析思路,主要是为了学生了解有这样一种类型的空间分析,即基于矢量地图的空间统计分析,并不要求学生去记忆其中的具体内容。

2. 基于地理空间和距离拓展概念所进行的空间分析

这种空间分析是在突破传统的欧氏(欧几里得)地理空间和距离观念的基础上,基于拓展的地理空间和距离概念所进行的空间分析。这方面的代表著作之一是 Distance and Space: A Geographical Perspective (A. C. Gatrell, 1983),我们在§5.2.3中已经接触到空间距离的多种含义,除欧氏距离外,还有时间距离、经济距离或成本距离、风险距离、认知距离、社会空间和生态距离。事实上,这些拓展的空间距离概念是与拓展的地理空间概念相应的。这些拓展的地理空间概念包括:

- 时间空间。一般而言,它基于地点间的旅行时间而建立。
- 经济空间。与经济地理研究关系密切,基于地点、地物间的经济关系而建立。
- 认知空间。基于人类认知中的地点间距离而建立。
- 社会空间。基于地点间的社会亲和程度而建立的空间。
- 生态空间。基于地点间的生态关系,即地点间的相互环境影响及其对它们所共享的环境之响应关系而建立。

上述每一种空间都有进一步的区分和研究内容，目的是较合理地确定构建该种空间所需的地点、地物间的相互关系值。例如，我们可以通过各种分析方法，给出某区域各城市两两间的社会距离，并据此构架社会空间；可以让学生写出他们所认为的各城市两两间的距离，并据此构架学生对此的认知空间。进而，根据这些城市两两间的社会距离或认知距离，可以可视化表现该空间的、相对原地图变形的形态。

能辅助进行这种研究的方法称为 MDS(multidimensional scaling)。MDS 是一种算法。用户应用 MDS 时，首先要按照某一种“空间”的概念，将研究地区地点两两间的距离或“差异”形成一个距离矩阵或“差异”矩阵，输入 MDS。MDS 将为用户寻求一个尽可能近似地表达输入差异的空间结构形象。

虽然基于地理空间和距离扩展观念的空间分析，目前仍处于研究探索阶段，但它试图从不同的角度，突破空间几何关系分析，探索某种基于地理空间位置的、但又含有非空间属性关联的空间关系，拓展了思路和视野，具有积极的研究意义。

3. 分析地图学简介

分析地图学是研究广义地图学内在的定量和数学关系的学科，它也试图发展现有空间理论的概念、范畴和应用。分析地图学与地图学的其他分支有很大不同，它是地图学、离散数学、地理学、计算机科学和图像分析的交叉科学。分析地图学与地理信息科学(GISci)有相当的重叠，它可能扩充地理信息科学的理论基础。

分析地图学基本概念和理论包括：

- 地图转换。空间坐标、空间相应类型和地图数据结构的转换等。
- 真实和虚拟地图。虚拟地图是不直接可见或不具备有形实体的地图，如屏幕上、存储器和数据库中的地图。真实地图与虚拟地图之转换的概念已被用于确定空间数据系统的逻辑结构和交互式空间数据系统设计中。
- Nyerges 数据层。它清楚描述了从现实世界中进行空间信息抽象，从真实地图向虚拟地图转换过程的 6 个层次：现实数据层、信息结构层、典型结构层、数据结构层、存储结构层和计算机编码层(可参照本书 §2.1.1 理解)。
- 空间原型对象。定义一系列 0 维、1 维、2 维的空间原型和简单对象，可用于构建几乎所有类型的、从 0 维到 3 维的数字空间数据对象，并可应用于空间数据科学的软件中。
- 抽样法则。合理利用这一基本的空间地图理论有利于进行合理的抽样、插值、邻接算法、像元分辨率之确定和许多其他应用。

分析地图学业已开展研究的较重要领域包括：空间频率分析、空间邻接算法、适合自由分布的确统计空间分析、傅立叶理论计算规则栅格地表 Z 值空间频率和振幅、空间分形算法、特征地形线的 Wartz 网络、不规则栅格分析、多边形叠加的算法研究、制图综合、空间对象形状分析、空间数据模型和结构、时态数据、地图数据质量、地图数据转换标准、地图投影和空间数据插值等。分析地图学目前的研究课题主要有：不规则栅格数据、真三维空间对象的构建研究、时空模型和分析可视化。

特别值得一提的是分析可视化(AV),它将可视化看做空间分析过程或结果的有机部分,而不仅是空间数据的表达。这一点也使分析可视化区别于虚拟现实模型语言(VRML)。AV研究还包含地形晕渲图与卫星影像之叠合及与全息摄影技术之结合、交互虚拟可视化环境和动态地图、透视晕渲的改进算法、通过空间可视化利用空间知识数据库的新方法等。

5.3.3 进一层的空间分析

本节介绍较复杂的,或进一层的GIS空间分析。

1. 多种基本空间操作或分析功能的组合

解决涉及空间分析的实际问题,通常需要应用多种空间分析操作或分析功能,将它们有机组合起来,实施空间分析过程。

例如,某道路拓宽改建,计划从原有的20m拓宽至60m,需要进行拆迁补偿金估算,即根据不同类别、等级房屋拆迁的补偿标准,估算拆迁补偿费的总值。

为此,首先需要进行目标分析。由于补偿标准使按面积计算的,要估算拆迁补偿费,先要通过空间分析计算该项工程将拆迁的各种不同类别、等级的房屋的总面积。

目标明确后,即可准备分析的数据:一是城市规划部门的1:500比例尺现状地形图;二是房管部门的有关房屋不同类别、等级的地图数据。

接下进行各种空间分析操作。首先选择拟拓宽的道路,根据拓宽半径,建立道路的缓冲区。然后将此缓冲区与现状地形图进行拓扑叠加。其次,在生成道路缓冲区的现状地形图上进行空间查询,即检索出全部或部分位于拓宽区内的所有建筑物。注意,部分位于缓冲区的房屋也要拆迁。对这一部分房屋的检索方法可以基于房屋原标识号,同标识号的房屋多边形只要有一个或以上多边形进入缓冲区(带有缓冲区标识),它们皆将为拆迁房屋。然后,再进行空间操作,即将拆迁房屋多边形全部选出,并剪下(clip);剪下的地图与房管部门的有关地图数据再进行拓扑叠加,形成新的地图。

再下一步,进行统计分析,即统计不同类别、等级的拆迁房屋的多边形面积;在此基础上,计入建筑物的楼层数因子,计算出不同类别、等级的房屋实际拆迁的总面积;根据房屋拆迁的补偿标准,计算拆迁补偿费的总计值。

最后,将分析结果以地图和表格的形式打印输出。

2. 应用领域的专业模型与GIS相结合

GIS可应用于数十种领域。因此,应用领域的专业模型常嵌入GIS系统,与GIS的空间操作与分析功能,以及数据库、可视化等功能相结合。虽然目前大多数GIS软件对专业领域的动态模拟尚缺乏能力,但是,现有水平的结合也是很有意义的。

例如,美国加州某重点生态保护区欲开发一新的商业中心,有三个可选地

址。项目的环境评价中要求计算每个新中心选址将引起的地面径流量(runoff)的增加,这种增加会导致水土流失。该 GIS 系统采用了一个水文模型:

$$单位面积径流量\ Q=(P-0.2S)^2/(P+0.8S),$$

$$其中,S=1\ 000/CN-10$$

式中,P 为 24 小时降水量;CN 是一个体现土壤渗透性能的因子。

CN 取决于三个空间因子:土壤类型、土地利用类型和前期湿度条件。前期湿度条件体现具体地块的平均湿润情况,分为几个等级,等级之值由当地 24 小时平均降水量确定。这样,GIS 系统要进行上述三层因子的空间叠置分析,查算出每个叠置单元的 CN 值(包括运用逻辑运算手段,参见 §5.2.2 4);进而通过上式求得各单元的单位面积径流量 Q,以及全区域总径流量。

上述空间叠置要进行两次。一次采用原土地利用图,计算无此建设项目情况下的全区域总径流量。另一次采用项目建设后变更的土地利用图,计算若按某种选址进行建设后将导致的全区域总径流量。如此可计算每种选址导致的总径流量的增加,用于选址比较。

显然,上例中空间操作分析与水文专业模型缺一不可,二者相结合,相当满意地解决了一个实际的环境问题。

GIS 与应用领域的专业模型的结合主要有三种方式:松散结合、交互界面和镶嵌结合。松散结合中,专业模型与 GIS 各自独立,通过格式转换相互调用结果。交互界面的方式开发支持模型的 GIS 数据库,同时开发统一界面,既可以为模型提供输入数据,又能利用模型结果进行处理和演示,所有的数据转换是通过交互界面自动进行的。更进一步的方式是镶嵌结合,这类结合一般用 GIS 的函数命令重写模型。三种方式各有利弊。

图 5-7 是一个水力/水文模型与 GIS 相结合的洪水预警模型实例,该结合接近于镶嵌结合。模型可给出浙江省分水江流域在任何降雨条件和前期湿度条件下(以该流域若干雨量站和蒸发站资料为输入数据),分水江 94 个河段或河道中任一个河段,在未来 108 小时内的水位和流量等预测,以及洪峰演进情况(图中曲线分 108 个格子)。

3. 空间区位-配置模型

企业和公共服务业的机构、设施之地理位置对其经济效益或自身发展具有至关重要的作用,因而需要对设施或设施网络的供给和需求之间的相互作用关系进行分析,进行机构设施位置的评价和优化。机构设施区位评价是对于现有服务设施的空间位置分布模式的评价,机构设施区位优化是对于其最佳位置的搜寻。在现实市场功能行为研究基础上,人们提出了不少能够针对现实世界的市场功能区域进行空间分析和模拟的模型,如克理斯塔勒的中心地理论、Reily 的零售重力模型和 Tobler 的价格场和作用风,以及众多的修正式等。这些模型

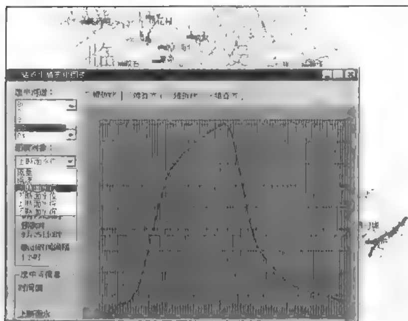


图5–7 浙江省分水江94河段未来108小时水位预测

为GIS的空间区位–配置模型提供了思路和一定的应用基础。

空间区位–配置模型用于解决位置–分配问题。位置–分配问题是在规划重要公共设施，如医院、幼儿园、游戏场所、养老院、学校、警察局、消防队、急救站和其他管理设施等的位置及其附属区域而产生的，并且可应用于企业、商业、服务分析。位置–分配问题一般包含供需两方面因素：需方是一定数量的居民集中点，这些点被称为需求点（或消费点、居民点）；供方是一定数量的供给点（某种公共设施）。位置–分配问题就是要对供需双方关系进行分析和规划。例如，已有需求点，求供给点，即选址或定位；或已设供给点，求分配，这是配置或确定供给点附属区域的问题；或者同时求供给点和分配，则涉及位置–分配或定位–配置问题。回顾§5.2.4的网络分析，不难理解上述概念。

区位–配置的优化模式基本结构由一系列边界条件和一个（或几个，但少见）目标函数组成：在这些边界条件下，求目标函数的极大值或极小值。边界条件代表了规划目标所必须满足的规划条件，代表了对于目标规划区域功能的基本评价；而优化目标函数（即求目标函数的极值）则代表了一般最大限度可能达到的规划目标。因此，在边界条件中体现出来的有关目标函数的规划条件具有首要意义。引入目标函数的极大、极小化意义，在于得到一个定位–配置问题的明

确答案。

在空间优化过程中，如果目标函数和边界条件都是线性的，则可采用线性规划方法来解决问题。所谓线性规划，是指在一组线性的等式和不等式的约束下，求一个线性函数的最大值或最小值的问题。欲进一步了解关于空间区位－配置信息模型，可参见陈述彭等的《地理信息系统导论》第4章第8节。

4. 专家系统

专家系统是人工智能在信息系统中的应用。人工智能的目的是用计算机来模拟人类（动作、视听能力、语言和人脑等）。在人脑模拟方面，目前主要是模拟领域专家的思维过程，因为领域专家有清楚的知识与逻辑。专家系统内部具有大量专家水平的某个领域知识与经验，它们能够用来解决该领域的问题。设计专家系统的关键是知识表达和知识运用。目前专家系统的知识表达主要基于"产生式规则"，即使用"If…Then…"语句。

一般的专家系统包括数据库，知识库，推理机，解释器及知识获取五个部分组成，它的结构如图5－8所示。

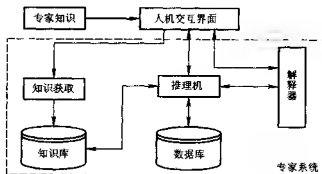


图5－8 专家系统结构

知识库用于存取和管理所获取的专家知识和经验，供推理机利用，具有知识存储、检索、编辑、增删、修改和扩充功能。

数据库用来存放系统推理过程中用到的控制信息，中间假设和中间结果。

推理机用于利用知识进行推理，求解专门问题，具有启发推理、算法推理；正向、反向或双向推理；串行或并行推理等功能。知识推理方法可分为图搜索方法和逻辑论证方法两类。

解释器用于作为专家系统与用户的人机接口，其功能是向用户解释系统的行为，将用户输入的提问及有关事实、数据和条件，转换为推理机可接收的信息，以及向用户输出推理的结论或答案，并且根据用户需要对推理过程进行解

释，给出结论的可信度估计。

知识获取器则是专家系统与专家的“界面”。该界面是知识工程师（专家系统的设计者）采用专题面谈和口语记录分析等方式获取知识，经过整理后，再输入知识库。为了提高知识工程师获得专家知识的效率，可以借助知识获取辅助工具来辅助专家整理知识或辅助扩充和修改数据库。近年来，开始机器学习、机器识别、半自动化等方法获取知识。

从上述可见，专家系统可广泛用于各种专业领域，地学、资源和环境的众多领域也不例外，业已出现了很多专家系统。这些专家系统本身都是独立的系统，但由于GIS系统与专家系统同属信息系统范畴，专家系统非常容易自然地嵌入GIS系统中，作为GIS系统上层建筑的一部分（参见图1－4）。

专家系统与GIS系统结合的紧密程度不尽相同。一般而言，专家系统与其所嵌入的GIS系统共用界面和关系数据库；但有的专家系统相对独立一些，有的则交织得多一些。后者例如，用于遥感分类的专家系统，就常与其他遥感分类功能相辅相成，发挥很好的作用。

5. 空间决策支持模型

本小节主要引述邬伦等的《地理信息系统》中的论述。

决策是一个决策者为达到某种目标或目标集合，根据一定的约束条件下在多种候选方案中进行选择的复杂过程。空间决策是有关空间行为的决策。GIS经常面临大量空间决策，如区域规划、土地利用规划、设施位置选择和环境管理等的决策问题。空间决策问题大大超过了地理信息系统通常的空间分析功能的要求。为此，空间决策支持通常要借助于计算机领域的决策支持系统和专家系统等的已有成果。

决策支持系统（DSS，decision support system）是辅助决策者通过数据、模型、知识以人机交互方式进行半结构化或非结构化决策的计算机应用系统。它是在管理信息系统（MIS，见§1.1.3.2）的基础上发展起来的、为解决结构化、非结构化和半结构化的问题提供了更广泛的方法。DSS为决策者提供分析问题、建立模型、模拟决策过程和方案的环境，调用各种信息资源和分析工具，帮助决策者提高决策水平和质量。决策支持系统是辅助管理者对半结构化问题的决策过程，支持而不是代替管理者的判断，提高决策的有效性而不是效率的计算机应用系统。

人类的知识可分为结构化和非结构化两种知识。结构化知识有着高度结构化的形式和结构化的求解程序，包括数学模型、统计方法和计算机算法等都属此类型的知识，它们在表现和分析方面遵循固定的框架，大多数情况下只能被专家理解，又称为程式性知识（procedural knowledge）。然而大量的知识都是非结构化的，像人类的体验、直觉、价值观和专家经验等，本质上是定性的，不能用固定的程序进行表达，又称为描述性知识（declarative knowl-

edge)。决策者使用信息和知识,在解决结构化、非结构化和半结构化问题上的复杂程度大不相同。以设施配置为例,在某些特定约束条件下确定设施配置的最少数量是一个结构化问题,可以通过最优化方法进行求解,寻找最优设施配置的所有可能的位置则是一个半结构化问题,涉及多种准则评价和价值评判;为设施配置确定总体目标和总体方针政策则属非结构化问题,涉及灵活的定性问题,不能用固定的程式性知识来解决。

DSS 的基本结构主要由四个部分组成,即数据部分、模型部分、推理机部分、人机交互部分,如图 5-9 所示。图中可见,决策支持系统包含了专家系统的成分。

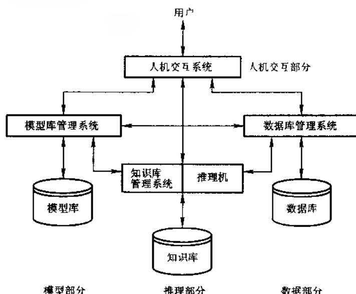


图 5-9　决策支持系统的组成

空间决策支持系统与一般的决策支持系统性质相同;但空间决策更注重空间数据和空间问题的获取和解决。由于空间数据和问题的特殊复杂性,空间决策支持是一个涉及多目标和多约束条件的复杂过程,通常不能简单地通过描述性知识或程式性知识进行解决,往往要求综合地使用信息和领域专家知识,以及二者之间的有效地交流。

由于空间决策支持系统的复杂性,目前的地理信息系统还难以达到空间决策支持系统的要求,尽管 GIS 本来的宗旨是面向空间决策。GIS 与 DSS 的不同之处在于其数据模型和数据结构的复杂性。地理信息系统为决策支持提供了强大的数据输入、存储、检索、显示的工具,但是在分析、模拟和智能推理方面的能力比较弱,其逻辑结构和智能层次不能满足解决复杂空间决策问题的需要,

特别是那些非结构化的问题。目前GIS还不适合用于对各种知识形式的处理，不能作为空间决策支持系统的神经中枢

所以，为解决复杂的空间决策问题，需要在地理信息系统的基础上开发通用智能空间决策支持系统，用于数据获取、输入、存储、分析和输出；用于知识表现和推理；用于自动学习、系统集成和人机交互。所用的新技术包括人工智能技术、知识获取、表现和推理等知识工程技术，以及集成数据库、模型、非结构化知识及智能用户界面的软件工程技术。

开发通用智能空间决策支持系统最经济和灵活的方式是应用软件工程和知识工程的方法，开发空间决策支持系统的开发环境或通用开发工具，使领域专家可以用它快速、高效地建立多种领域空间决策支持系统。开发环境的核心是一个专家系统壳(Shell)，它可以单独作为专家系统开发工具，直接控制SDSS的控制流和信息流，提供表达和存储非结构化领域知识；它还包含推理控制、系统和用户界面对外交流的元知识，以及非结构化空间知识的推理机。它是SDSS的大脑。为使用空间和非空间数据，专家系统壳有一个与外部(GIS、关系数据库和遥感信息系统)的接口。SDSS还应包含一个模型管理系统，管理和处理程式性知识包括算法、统计程序和数学模型；它也有接口通向专家系统壳，可以通过专家系统壳的元知识进行调用。除了与数据库管理系统、模型管理系统的接口外，友好的用户界面和知识获取模块也是专家系统壳的基本组成部分。

第六章
地图制图及输出

地理空间数据在地理信息系统中经过处理、分析，所提取信息必须以某种可感知的形式输出，供有关用户应用。矢量数据输出和栅格数据输出的基本原理已在第二章中介绍。本章要讲述的，则是地理信息系统的输出产品，主要是地图产品及其绘制；而地图绘制的主要内容是地图符号及注记制作。

本章第一节介绍 GIS 输出和 GIS 地图绘制的一般概念；第二节介绍 GIS 系统中地图符号的具体制作方法；第三节介绍地图注记、排版和最后输出。

§6.1 地图绘制与空间对象的符号化

6.1.1 GIS 输出与地图绘制

1. 地理信息系统输出产品

地理信息系统的输出是指经由系统的分析和处理，为从事研究、规划和决策的人员提供可以直接使用的产品，它们是地理信息系统所输出信息的表现形式，反映 GIS 所研究、处理的地理实体的空间特征或属性特征。地理信息系统的输出的主要形式有地图、图像、统计图表、文字和多媒体等。

- 地图。只要不特别说明，地图产品皆是指矢量数据形式的地图。地图是空间实体的符号化模型，它遵循一定的数学法则，将现实世界所传递的地理信息通过概括总结，并运用特定的符号表示在一定载体（如纸张、计算机显示器）上，用以传递地理空间中地物实体集合的质量、数量和时空特征及其发展变化的规律。

- 影像地图等栅格形式的 GIS 输出产品。栅格图像形式的 GIS 输出产品有多种。栅格图像通过画面每个细部的明暗或色调来表现周围实体的界限和形态，具有丰富的表现力。因此，经过一定整饰的卫星影像和数字航测图像等可以直接作为 GIS 产品输出。例如，经拼接或剪裁的地区或城市的卫星图像或航测图像，常作为办公室悬挂物，直观而形象地表现区域的地理空间形态。经

较多加工的现代栅格形式的 GIS 输出产品主要有下述几类：

一是经过分类等遥感图像处理的卫星分类图像（§ 4.1.2.3）

二是经过三维处理的透视立体图像，如图 4－19、4－20 和图 5－1。图 5－1 实际上是动态漫游中的三维虚拟现实动画图像中的一个画面。有些透视立体图像上还进一步叠加上地图信息，形成具有符号化点、线、面空间对象信息的三维透视立体图像。

三是影像地图。影像地图是经过正射校正，即由多中心投影或中心投影转换为垂直投影的卫星影像和数字航测图像，与矢量（线划）地图叠加的产物。这里，矢量地图是在栅格化后，被叠加到图像上的。影像地图既是地图，具有符号化的点、线、面地物信息，又有栅格图像的丰富表现力，是一种比较新型的地图输出产品。

• 统计图。统计图利用图或图形来表示线实体的属性特征及属性间相互关系，由于采用可视化手段，统计图使用户对这些信息有全面、直观的了解。统计图常用的形式有柱状图、扇形图、直方图、折线图和散点图等。统计图常被结合到专题地图输出中，表现点状和面状地物（有时也表现线状地物）的属性统计特征、属性间相互关系或数量的变化。例如，在某省环境保护图中，为表现该省几种主要污染物在每个地区（市）的比重，可在每个地区（市）中给出表现几种污染物的柱状图或扇形图（参见图 6－14）。

• 文字和表格。地理信息系统有时也输出单纯的文字、数字或表格。例如，从土地管理系统中输出一段有关的土地法令、法规的文件；输出不同分区土地价格的统计表等。

• 多媒体。现代很多地理信息系统都具有多媒体功能，带有录相和录音等资料。例如，在水利 GIS 系统的屏幕地图上用鼠标点击某水库的符号，便可以播放该水库的录相资料；GIS 系统的动态三维虚拟现实图像也可以作为动画文件或电影文件输出等。因此，现代 GIS 系统的输出还包括多媒体视频、声频等的输出。

在上述地理信息系统的输出产品中，最常用而重要的产品是地图。

2. 地图绘制、GIS 输出与 GIS 可视化

在地图绘制、GIS 输出与 GIS 可视化的关系问题上，有关 GIS 书籍的作者存在一些不同的提法和理解。我们认为：

第一，地理信息系统输出产品并非都是可视化的；但可视化输出是地理信息系统输出产品的主流形式。

从上述地理信息系统输出产品的介绍可见，单纯的文字、表格和声频等 GIS 输出都不是可视化形式的输出，但地图、影像和统计图等大多数重要的 GIS 输出产品是可视化的，它们是地理信息系统输出产品的主流形式。

第二,可视化输出并非地理信息系统可视化内容的全部,但它是地理信息系统可视化处理或功能中,非常重要的层次提高的一部分

地理信息的可视化贯穿于GIS整个系统的功能和流程,从数字化输入、数据结构的构架、数据编辑、数据处理和分析等,直至输出,都浸透着可视化表达的理念,离不开人机互动的可视化的环境。但是,同GIS其他部分相比,地理信息输出的可视化层次更高。这是因为,在矢量GIS系统的可视化环境下进行数据编辑、处理和分析的点、线和多边形,是没有符号化或没有完全符号化的,它们能为GIS系统的操作人员所理解,其属性能随时查询;但是,要为从事研究规划和决策的用户所应用,它们就必须进一步符号化,将用户需要的地理信息用直观形象的符号表示出来。另一方面,从栅格数据输出的角度,影像地图和三维透视立体图等,不言而喻,也是层次较高的可视化输出。

第三,地图是可视化的GIS输出的最重要方式。

在地图、影像和统计图等可视化输出产品中,地图(包括与图像叠加的地图产品)又是最主要的输出产品。地图绘制是空间对象的符号化过程,经过符号化的地图具有可量测性、直观性、精练性和综合信息一览性等优点,是反映地理事物和现象的时空分布、组合和联系,揭示其发展变化的最有效手段,因而至今仍是应用最广泛的GIS主流输出产品。

下面,我们便将学习的重心转到矢量地图绘制及输出上。

3. 地图绘制:空间对象的符号化过程

在矢量GIS中,点、线和多边形空间实体用位置数据和属性数据来表示,如果不符号化,输出的只是几何的点、线和多边形(几何点有时甚至看不见)。因此,矢量GIS输出,无论是在屏幕上显示还是硬拷贝输出,必须在点、线或面状地物坐标规定的位置上,绘制表现该地物某些地理信息的符号或文字。空间对象的这种符号化过程,就是空间实体在地图中以各种地图符号表示出来的过程,是地图绘制的主要内容。

地图符号是地图的语言,包括图形符号和地图注记。图形符号由形状不同、大小不一、色彩有别的图形组成;地图注记是地图上文字和数字的总称。地图符号用以直观形象、一目了然地表达空间对象的位置、形状、质量、数量特征和相互关系,以及区域总体特征(图6-1)。地图符号的运用是一个人类视觉感受与认知的过程。

按所表达的空间对象的不同,地图符号分为点位符号、线状符号和面状符号三种主要类别。点位符号绘制在点状实体处或面状实体的中心,采用点状图形表示点状或面状空间实体的特征;线状符号采用线状图形表示线状空间实体的特征;面状符号在面状区域内一般用填充模式表示区域的特征。

无论是点位、线状和面状符号,在地图上大多有确定的位置,一般不能随便

地图符号	图形符号		地图注记
点位符号	●	◎	
线状符号		—·—·—·—	北京 12ab
面状符号	[斜线填充]	[波纹填充]	

图6-1 地图符号

移动(但当作为说明符号时则不一定要求精确定位)。地图符号可以采用正射投影、透视投影和会形、会意的方式来绘制。按比例关系,地图符号则可分为三类:依比例符号、不依比例符号和半依比例符号。

关于地图符号系统及其绘制的丰富内容,可参见任一种地图学教材。本课程讲授地图绘制的重点,是GIS系统中的地图绘制及输出的原理和特点。

有些地图学者将地图注记与地图符号看作是不同类别的地图语言,另一些地图学者将地图注记看作是地图符号系统的一部分,本书采取后一种观点。此外请注意,一般地图学中的"点位符号"不描述面状实体而只描述点状实体,与上面给出的定义有所差异。这是因为,与传统地图不同,很多矢量GIS中的多边形有标签点(§2.2.2.2),在标签点处给出描述多边形属性的点位符号,是方便而有效的。

4. GIS中的空间实体符号化过程的一般原理和步骤

在GIS系统中实施空间实体符号化,首先需要建立点位、线状和面状符号的基本符号库。在所有GIS软件平台中,都有现成的基本符号库,并允许用户在基本符号库的基础上进行地图符号的再制作,即符号定制。用户再制作的地图符号可以用于现场的地图绘制,也可用以充实基本符号库,即作为新品种添加到基本符号库中。基本符号库中,每一种点位、线状和面状符号都有符号代码或符号索引。例如,图6-3显示线状符号库的一部分,每种线状符号的编号即符号索引,如代码为77的线状符号为一种铁路线符号等。

在基本符号库的基础上,GIS中的地图符号化过程一般经历如下步骤:

- 获取需要符号化的空间实体的几何坐标和属性参数;
- 巡查基本符号库:如果没有输出所需的地图符号,则在基本符号库的基础上进行地图符号的再制作,即符号定制
- 按照地图绘制及输出的要求建立各属性数据代码与所选定的输出地图符号的代码之间的对照表(lookup table)

例如,属性为61的线状空间对象为双轨铁路,而符号库中(图6-3)代码为69至77的线状符号都可以是铁路线符号;假定用户从中选择代码为77的线状

符号作为最后输出的铁路线符号，那么，在对照表中就将属性代码 61 与符号代码 77 对应。

● 根据符号代码到符号库中获取符号描述信息，并根据空间对象的符号描述信息和几何位置数据对空间实体进行符号化，并将符号化结果输出。

例如，根据对照表中属性代码 61 与符号代码 77 的对应关系，GIS 系统便到符号库中获取代码为 77 的符号描述信息，以及属性代码为 61 的线状地物的几何位置数据，然后在输出时，在属性代码为 61 的线状地物的几何位置上，全部绘制符号代码为 77 的铁路线符号。

在上述第三步，即根据地物属性绘制指定符号的方面，不同 GIS 系统处理方法还不尽相同。一般而言，上面所说的通过对照表，建立空间实体与符号宽松联系的方法是比较灵活的，它不仅允许随时改变地物的显示符号，多种地物可以选择同一符号输出，不必重复设计相同的符号，而且可在地物的多重属性中，灵活地选择某种属性来控制地图符号的绘制及输出。有些系统缺乏这样的灵活性。例如，某些 GIS 系统要求符号代码与地物编码严格一一对应，符号库中，一种地物编码对应一种符号。

下面对 GIS 中的点位、线状和面状符号进行某种程度的具体描述，以便使读者对这些符号有一些具体的印象。

6.1.2 地图符号及其分类

1. 点位符号

点位符号表达可以在地图上忽略其实际大小或可视为点状的小面积地物，如城市、车站、山峰、测量控制点等，也可用来表达面状地物的某些属性特征。点位符号绘制在点状实体处或面状实体的中心。点状符号，可以是空心的或实心的点、圆、多边形等图形元素，也可以是多种图形元素的组合（图 6－2）。实心点内部经填充的点，其填充方式是可以是一般填充或位图填充。点状符号有以下特征：

第一，点符号的图形固定，一般不随图幅大小及其位置的变化而变化。

第二，点符号一般有确定的定位点和方向性。

第三，点符号图形大都比较规则，由几何图形构成，简单、美观、形象，易用数学公式表示。

在 GIS 中制作图形符号时，点符号常常是绘制线、面符号的基础，是线、面符号的基本组成部分。下面我们将看到，线符号中往往包含点符号，面符号中包含线符号和点符号，如面符号的边界线为线符号，内部则为按一定规则排列的点符号组合填充。

符号类型		示例
简单图形	弧线符号	
	折线符号	
结构图形	个体结构符号	
	组合结构符号	
扩展图形	长度扩展符号	
	面积扩展符号	
	容积扩展符号	
统计图形	直方图	
	曲线图	

图 6-2 点位符号类型

2. 线状符号

线状符号是表达线状地物，如河流、道路、管道线路等在地图上呈线性伸展的物体或制图现象的符号，这些线状地物的实际长度须依比例尺表示，而宽度不用按比例尺表示，属半依比例符号。线状符号有以下特点：

第一，线状符号都有一条有形或无形的定位线，即绘制符号的线状地物的 (x,y) 坐标串所决定的定位线。

第二，线状符号可以进一步划分为曲线、直线、虚线和平行线，以及沿定位线连续配置点符号等。

第三，不少线状符号可以进一步分解成具有单一特征的线状符号，即一线状符号可以由若干条具有单一特征的线状符号组成。

图 6-3 为 MapInfo 中常用的线状符号

根据上述特点可知，GIS 中线符号可以采用两种方法绘制：一是重复配置点符号图元法，即将线符号分解成基本点符号图元，然后沿着线符号的定位线，连



图6.3 线符号

续配置点符号;二是组合绘制方法,即任何线符号由单一特征的线符号组合而成,针对不同的线符号设计各种单一线型。线符号绘制时要考虑线的偏移量、线宽等;虚线、双虚线还要设置虚线之间的间隔、起始位置、线宽、实部长、虚部长等;重复配置点符号时,还要考虑点状符号的缩放系数、绘制端即起止点是否绘符号、方位符号是始终朝北还是指向转折点角平分线方向等。

3. 面状符号

面状符号是指在平面图上各方向都须依比例尺表示的符号,通常用来表示呈面状分布的空间对象,诸如湖泊、植被覆盖、土壤类型区等地物或地理现象。面状符号有以下特点:

第一,有一条有形或无形的封闭的轮廓线定位,即绘制符号的多边形地物边界的(x,y)坐标串所决定的闭合定位线。

第二,面状符号一般采用填充模式表示区域的特征,即在轮廓线范围内配置点状符号的排列、绘晕线或普着染颜色等。

GIS 中常用的面符号绘制方法有三种:一是以晕线填充图案绘图区域,二是以按一定规则排列的点符号图案填充绘图区域;三是位图填充图案。晕线填充图案时要考虑填充线的倾角、线宽、颜色以及填充的起始位置、偏移量和实部虚部的长度。点符号填充图案则要设计填充点符号的行列偏移、行列间距、缩放

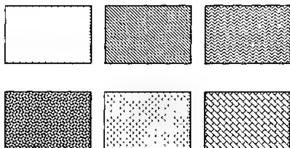


图6-4 几种填充图案

系数以及是固定还是随机旋转角形式、是品字形还是井字形存点形式。而位图填充图案要考虑位图的长宽、缩放系数、旋转角、填充形式(品字形、井字形)等。

6.1.3 专题地图制图特点

1. 普通地图和专题地图

地图有不同的分类方法。最普遍的分类是按所反映的内容分为普通地图和专题地图两大类。普通地图是以相对均衡的详细程度表示区域内各种自然合社会经济现象的地图,其基本内容有水系、地貌、土质植被、居民地、交通线和境界等六大地理要素,以及测量控制点、独立地物、管线和垣栅等要素。普通地图主要包括地形图、地理图与平面图三类。地形图是最常用、最典型的普通地图,为国家各项建设的规划设计与施工、军事指挥和科学研究的参考用图,是其他地图绘制的基础底图。地理图是以高度概括形式反映广大制图区域内最主要的地理要素和区域重要特征的地图。平面图是指在小范围内只表示地面要素及平面位置而不表示地表起伏形态的地图。普通地图多由专业测绘部门和地图学家绘制。

专题地图是以普通地图为地理基础,着重表示制图区域内某一种或几种自然或社会经济现象的地图。专题地图可分为自然地图、社会经济地图、环境地图和其他地图。

普通地图一般由专业的测绘部门和地图学家所绘制。而对大多数地图用户,包括地理信息系统工作者和地学、资源、环境和区域管理等领域的GIS用户来说,地图绘制主要是专题地图的绘制,他们通常在地形图的基础上,着重表达、分析和输出自己的领域的专题内容。目前在这些领域内,专题地图的应用十分广泛,深入不同的主题。尤其是在自然资源和社会经济的调查研究中,应用专题地图能更有效地分析和研究问题,专题地图的编制获得了迅速发展。

因此,有必要简单讨论一下专题地图绘制的特点。事实上,普通地图和专题地图的区别,关键不在形式,而在内容。在表现形式上,它们可能都采用某些表现方法、符号和颜色;但是由于表现内容的多样性,专题地图绘制在属性数据的符号化表示方面,还是形成了自己的一些特征,专题地图的地图符号在表现多种类别的专题要素时形成了很多的表现方法。为此,首先简单介绍一下不同类别的专题地图。

2. 专题地图分类

专题地图种类繁多,目前还没有统一的分类标准。现在通常从下述几个角度对专题地图进行归类。

• 按地图内容划分。按地图内容划分,也就是从学科的角度来划分专题地图。由于专题地图着重表示的自然和社会经济现象通常对应于不同的学科,按内容或学科角度分类是最普遍采用的一种专题地图分类方法。专题地图按内容可以划分为自然和社会经济两大类别。

自然地图主要表示自然要素的分布特征。自然地图又可以按学科进一步分为地质图、地貌图、土壤图、植被图、气候图、水文图、地球物理图、动物地理图和景观类型图等。这些专题地图还可以再进一步划分。例如,土壤图又可以再划分为土壤类型图、土壤化学成分图和土壤侵蚀图等等;气候图可以再划分为太阳辐射图、温度图和降雨量分布图等。事实上,随着自然科学研究的不断深化,自然地图按学科分类也愈来愈细,能更具体地反映不同类型专题地图的主题内容。

社会经济地图反映人类社会经济现象的地理规律。社会经济地图包括农业地图、工业地图和交通运输等产业地图,以及人口地图、行政区划地图、文化教育地图和历史地图等。社会经济地图通常也包括一部分自然要素的内容,因为人类生产活动离不开自然条件和自然资源的开发利用。例如,农业地图可以包括农业自然条件图、土壤资源利用和评价图,以及农业经济图等等。

• 按地图服务对象和编制目的划分。内容相同的专题地图常因服务对象和编制目的不同,而在比例尺选取、表示方法、内容取舍和概括程度等方面有不同的要求。例如,为科学研究编制的地貌图着重反映地貌形态成因和类型;而为农业服务的地貌图则详细反映地面坡度和切割密度等实用性指标。又如,表达地形的地图若为工程设计服务,就必须采用等高线精确描绘;而若从教学和旅游出发,则可用地貌晕渲法表示地形的立体形态等。

因此,专题地图又常从目的和意义的角度分类,例如科学参考图、实用工程图、教学图和宣传普及图等等。由于专题地图具有广泛的服务对象,各种意义的专题地图还可以再划分为若干类型。如科学考察地图可区分出科学一览图,部门参考图和预测预报图;实用工程图可划分为现状图、规划图和改造措施图;

宣传普及图又可区分旅游图和政治宣传图等。

此外，还有为专门目的服务的专题地图，亦称为专用地图，如航海图、宇航图和油气资源评价图等，它们具有特殊的用途、编制特点和制作方法。

• 按比例尺划分。同其他地图一样，专题地图也可按比例尺划分为大、中、小三种类型。通常认为比例尺大于1:10万为大比例尺专题地图；比例尺在1:10万至1:100万之间为中比例尺专题地图；小比例尺专题地图的比例尺小于1:100万。专题地图比例尺的大小关系到资料选取、内容综合、图形设计、使用方法等。例如，大比例尺专题地图往往要通过实地调查、野外测图或像片调绘获取专门内容，多为实测图；而中、小比例尺专题地图多数在室内编制或基于卫星影像判读等手段而制作。

3. 专题地图内容表示方法

由于专题地图的众多类别和专题地图表现内容的丰富多样性，专题地图绘制在属性数据的符号化表示方面，形成了很多的表现方法，专题地图的点状符号、线状符号和面状符号也更加多样化。例如，专题地图中的两种面状符号，即连续面状符号和离散面状符号都有丰富多彩的内容。连续面状符号可表现具有渐变趋势的连续面状分布，如地形、地磁场、气压和温度等；离散面状符号呈散点状分布于整个制图区域内，可表现耕地、森林和沙漠戈壁等。除精确定位的资料，专题地图还有许多的统计资料，在地图上绘制统计信息也是专题地图的特色。普通地图一般不表现统计资料。

专题地图多样化的点状、线状和面状符号和统计资料表达，构成专题地图丰富的表现手段，基于这些手段，不同内容、资料来源、服务对象和编制目的专题地图得以采用不同的表示方法以达到特殊有效的结果。

通常的专题地图表示方法有十余种。下面我们将其归纳为四种大类别来加以叙述：

• 符号法。符号法可分为点状符号法、线状符号法和动线法。

点状符号主要用于有精确定位的点状地物表示。专题地图上采用的点状符号按其形状可以划分为文字符号、几何符号、特征符号和艺术符号。符号大小和分级，可采用按连续绝对增长的比例或分段增长的比例等原则。符号的不同颜色和线划反映制图对象的内部结构。例如，用圆形符号表示城市时，圆的大小可表示城市的规模，圆内线划反映城市的行政级别等。风玫瑰图可以表示多种指标，如用圆符号定位于观测点，用不同方向的齿长表示风向的频率等。

线状符号法用于表示具有线性分布特征的制图要素，如海岸线、地质构造线等。线状符号采取不同的颜色、线宽和线型等。地物的类型可以用线型和颜色加以区分；地物的数量特征则可用符号大小或线宽表示。例如在交通运输图上，用条带宽来表示运输量，同时条带内用各种晕线或颜色区分货物品种等。

动线法相当于一种扩展的线状符号法，它用带箭头的线状符号（箭形符号）显示自然现象和社会经济现象的运动趋向及其数量、质量特征，如洋流和台风的运动轨迹、航程和旅游路线等。绘制箭形符号时，除根据表现对象选定线的颜色、线宽和线型外，绘制好不同形状、大小的箭头也很重要。

应用符号法编制专题地图的关键，是符号系统的选择和设计。一般要求符号系统的设计既能反映制图对象的内在联系，又要符合人们的视觉习惯；符号图形既要简明，又要能反映所表达对象足够多的信息。在专题图中，点状符号一般有精确的定位信息，当然符号的颜色和大小同时能传递更多的如数量等的信息。

• 等值线法。面状分布的图形要素常用等值线法来表示其连续分布或逐渐变化的数量特征。在制图时参照某一指标，把具有相等数量指标值的离散点相连成连续的线，这些线就是该指标下的等值线。实际地理现象中并没有这种标志，等值线只是一种表达整个制图地区特征的方法，反映制图对象的差异变化。等值线法还可采用灰度和颜色表示不同的数值或特性层次，明显地反映出制图区域内现象的分布规律，如地形高程、气压和降雨量等要素的大小强弱变化等。一般地说，等值线表示的是制图对象的绝对数量。而在等值线法表示逐渐过渡的统计数量（如人口分布）时，则使用相对值。相对值是指单位面积内现象数量的多少，如人口密度、开垦程度。相对值指标表示的是制图区域内的相对特性。

• 质底法和范围法。质底法和范围法属于定性表示面状要素的非统计方法。

质底法是为了反映制图区域连续分布现象的质量特征，把性质有差异的不同类型对象在地图上用不同颜色加以区分的方法。质底法反映整个制图区域质和量的差异，按确定的原则划分各种类型。质底法除了使用各种颜色外，还经常采用面状符号、晕线等表示手法。质底法应用的关键是划分图例类型和确定界线。在有些专题图，如地质图、土壤图和遥感植被图上，类型界线都是通过实地调查或像片判读确立的精确的界线。而有些专题图，如动植物地理分布图或自然区划图等，界线难于精确界定，属概略界线。图例分类系统和质量特征的说明是一个复杂的问题，常需要采用多指标分类方法，且与专业要求、区域研究程度以及其他相关分类技术密切相关。

范围法又称区域法，它按照特定现象要求，用颜色、线条、闭合的界线以及符号、文字等表示手段来反映该制图现象间断而成片分布的范围，例如煤田、油田、旅游区等。范围法往往用以表示在其他制图地区不经常出现的自然和社会经济现象。按制图对象分布特征在界线内出现的可能程度，范围法又分绝对分布和相对分布两种。绝对分布的界线是精确的，相对分布是概略的。二者绘制

的依据都是实测调查资料 对于难于精确反映分布范围的现象,往往用没有外围界限的一片点符号或阴影线等符号表示 范围法主要是表示制图现象质量特征,以定性为主,但有时也用颜色和晕线密度来反映分布强度的数量指标。

• 基于统计资料的方法。基于统计资料的方法有点描法、分级统计图法和图表统计图法等。

点描法基于使用统计,在制图现象分布的范围内,用大小相同的点群表示其分布特征 各区域单元内地理现象点的数量根据区域单元的统计数据计算。点数的密集程度对应于该现象在表示范围内的数量或发展程度,点越密该现象越集中,反之亦然 每个点代表的数量叫点的权值或称点数值。在出版的专题图上,点的大小多采用直径0.25～0.5 mm。一般而言,所采用的点愈大,所代表数值愈大;但点的数目会减少。

分级统计图法又称为分级比值法;图表统计图法亦称为分区图表法。这是两种反映统计资料分布特性的表示方法。分级统计图法用颜色深浅或符号疏密表示制图现象统计差异;图表统计图法是用柱状图、扇形图、直方图、折线图、散点图或其他图形符号表示各区域单元的统计数据。这两种方法的差别在于:用分级统计图法反映制图现象分布强弱,必须使用相对指标值;而图表统计图法则直接采用绝对值指标,通过图形符号大小对比反映区域差异。图表统计图法的符号不表示地物分布具体位置,只反映该区域单元的数量总和。例如经济图上经常使用圆形符号表示各统计单元的工业总产值,圆内按不同比例划分成小块来表示各工业部门所占的比重。

专题地图还有其他表现方法,就不一一叙述了。最后只提一点:很多专题地图的内容非常丰富,会牵涉到多种类型地物或信息的表示。因此,在实际的专题地图绘制作过程中,往往会在同一幅图中综合应用以上提到的多种表示方法。例如,定点符号法常与分级统计图法或图表统计图联合应用;线状符号法常与动线法联合应用等。

4. 专题地图的 GIS 实现

在地理信息系统中,专题地图的绘制实际上就是对空间对象的属性数据进行符号表达。在专题图的制作过程中,首先根据图层选择属性表和属性项,然后选择制图方法。

一般而言,对于进行定性描述的字符型属性数据,常采用符号法、质底法和范围法。质底法和范围法在地理信息系统中常称为"面状充填",系统会根据面状空间对象的特性在该对象范围内充填相应的颜色或面状符号或晕线等;范围的边界是否显示也很容易控制。这种过程与普通地图的面状地物的符号化过程基本类似。

对于进行定量描述的数值型属性数据,可采用符号法、等值线法、点描法和

统计图法等 数值型属性数据的专题地图绘制，比字符型属性数据要相对复杂一些。这时需要将属性值的数值大小转化为符号的大小。例如，若采用直方图表示工业产值，则需要将工业产值转换成直方图的高度，再标注在相应的标识点上 有时为了反映空间对象的某些指标值随时间变化的分布规律，需要将多个属性值制作成专题图符号放置在一个专题图上

§6.2　地图符号制作或绘制

上一节介绍了地图绘制的一些基本概念，本节讨论地理信息系统中地图符号的制作或绘制问题。GIS 系统中的地图符号绘制包括两方面内容：一是 GIS 软件的地图符号制作和符号库的建立；二是用户在 GIS 软件平台的符号库、符号工具基础上的地图符号再制作。

6.2.1 地图符号制作方法

这里所讲的地图符号制作，不涉及与硬件相关的具体操作方法，如绘图仪器如何抬笔、落笔或喷墨嘴如何动作等，而是从 GIS 系统在软件设计的过程中，如何有效、方便地实现符号绘制的技术思路角度着手陈述。

地图符号制作的方法很多，但可归纳为两种基本类别，即编程法和信息法。信息法又可细分为直接信息法和间接信息法两种。

1. 编程法

编程法是由程序员按照图形符号绘制的基本原理和要求，设计地图符号的绘图子程序，该绘图子程序能利用符号图形所需参数计算出绘图矢量，并能操作绘图仪绘制地图符号。一般而言，同一种地图符号或同一类地图符号编写一个绘图子程序。在绘图时按符号的编码调用相应的绘图子程序。子程序在调用时必须输入必要的参数，程序便根据已知数据和参数计算出绘图矢量并产生绘图指令，从而完成地图符号的绘制。绘制众多地图符号的许多子程序构成一个程序库。

编程法比较适用于较规则或可以用数学表达式描述的地图符号制作。例如，欲绘制类四边形，如常规矩形、梯形、正方形或圆角矩形等（参见图 6－5），可编写一个 Rectangle（矩形）子程序，只要给出起始点和终止点坐标或边角点的弧度等，则调用该子程序就能方便地绘制所要求的类四边形符号。编程法的缺点是增加符号不方便，要增加符号就必须修改和重编译程序库。对于非程序开发员的一般用户来说，灵活性和易用性较差；作为商业软件，则可流通性比较差。因此，目前这种方法的应用越来越少。



图6-5　编程法绘制符号

2. 直接信息法

随着当今可视化窗口系统和近傻瓜式应用要求的盛行，软件系统必须为用户解决尽可能多的基础算法或编程工作，使用户应用更方便，更有目的性。信息法就是顺应这种发展而产生的。信息法也称为符号库法，符号库描述符号的参数集合称为信息块。现在通用的GIS平台都采用符号库方法来进行地图符号的绘制，用户绘图时只要通过程序处理和调用已存在符号库中的信息块，即可完成符号的绘制。信息块内容随符号在地图上的表现形式不同而不同。一般符号信息块的构成有两种：即直接信息法和间接信息法。本小节先介绍直接信息法，下一小节再讲间接信息法。

直接信息法直接存贮符号图形特征点的坐标或具有一定分辨率的点阵，即常说的矢量数据和栅格数据信息，从而能直接表示各图形的细节特征的方法。这种方法查找和调用符号信息的算法相对简单，面向图形特征点而与符号图形无关，其思路如图6-6所示。

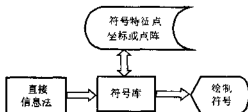


图6-6　直接信息法绘制符号

所谓特征点，就是表达某一图形符号所需要特别标注的点，它对描述符号细节有关键意义，如符号连线的拐点和接点等。在绘制地图符号时，绘图程序直接将符号特征点的位置信息转换为绘制地图的坐标（关于地图输出时计算机内的电子地图与输出设备的坐标及单位的关系，参见§4.2.3.3），然后将有序连接各特征点或输出点阵，绘成地图符号。该方法增加符号点比较方便。其缺点是获得符号信息较为困难，须静态占用存贮空间，特别当符号精度要求较高

时存贮空间的开销很大，实际调用比较困难，而且对符号进行放大操作时符号容易变形（如出现马赛克等）。

3. 间接信息法

间接信息法不直接记录符号点的坐标，而直接存放的是图形的几何参数，如图形的长、宽、间距、半径、方向角等，其余数据则在绘制符号时由绘图子程序按相应的算法计算出来。该方法的优点是占用机器的存贮空间小，能灵活表达复杂图形且绘图精度高，对符号进行无级放大时（指可以进行任意倍数的放大）不变形。图形绘制时所需的符号参数（如符号的ID、符号名）可通过系统交互获得。这里注意，符号参数和几何参数是两个不同概念。在绘制地图符号时，绘图程序必须先将符号图形的几何参数转换为用电子地图的坐标值表示，然后再转换为绘图输出的坐标系中的矢量数据（图形特征点坐标）或栅格数据，然后有序连接各特征点或输出点阵，绘成地图符号（参见图6-7）。

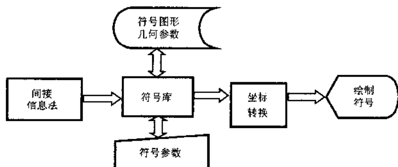


图6-7 间接信息绘制符号法

间接信息法的编程相对复杂，程序量较大，机器在图形绘制时需进行大量的运算，而且要比直接信息法多一个几何参数的坐标转换过程。但是，一旦软件开发商做好这一工作，用户使用就比较灵活、方便。因此，现今国内外的GIS软件中，大多采用间接信息块方法来绘制符号，并提供相应的符号设计模块。

6.2.2 矢量地图点、线、面符号的具体绘制

前面提到的地图符号三种制作方法是从系统设计的角度，介绍系统实现地图符号制作的概略思路，以便用户更好地理解GIS中符号系统运作的机制。本节介绍矢量地图点、线、面符号的具体制作思路和技术方法。

1. 矢量点符号绘制

在地理信息系统的地图绘制中，矢量点符号绘制占有重要的位置，这是因

为，不仅地图本身有很多点状地物需要表现，而且我们将看到，矢量点符号是线符号和面状填充符号绘制的基础之一，不少线符号和面状填充符号可以由重复配置的点符号组成。

从上一节图6－2可见，点符号实际上是一些规则图形，例如线段、折线、样条曲线、圆、圆弧、椭圆、特定三角形、矩形和四边形，以及其他多边形等，或者是这些图形的组合，这些图形作为点符号，存贮在符号库中。

当用户实际应用基本点符号绘图时，用户可以给出点符号绘制的基本参数。一般而言，基本参数有：定位点坐标（x_0,y_0）、缩放比例系数 λ 和旋转角度 θ 等。此外，GIS 软件还必须解决坐标自动转换的问题。地图符号库中保存的符号描述信息都是基于符号坐标系的，GIS 软件平台还必须根据用户的要求、电子地图和输出设备的坐标及单位，在绘制点符号时进行一系列缩放、旋转和平移等变换，将符号的矢量点坐标从符号坐标系转换到绘制地图的实际坐标系中的相应位置上来。

2. 矢量线符号绘制

线符号的绘制方法有两种：一是重复配置点符号图元法，即将线符号分解成基本点符号图元，然后沿线符号定位线连续配置点符号。这种方法的特点是每配置一个符号就要进行一次符号变换，变换速度随定位线的弯曲和符号的复杂程度而异，因而绘制速度较慢，且变换模板较难设计。二是组合绘制方法，它认为任何线符号可以由具有单一特征的线段单元（或基本线型）组合而成。例如，地形图上的线符号可归结为实线、虚线、点虚线、双虚线、双实线、连续点符号、齿线符号、带状晕线等13种基本线型，或者其组合。这种方法的特点是绘制速度快、算法相对简单，但要针对不同的线符号设计好各种线段单元的线型。

在 GIS 软件的地图符号库中，都有一系列基本线型。图6－3所示就是一部分基本线型。基本线型看上去都是直线形态的，但实际上不然；符号库存贮的基本线型，基本上都是关于其组成单元（点符号或线段单元）及其组合的有关信息。例如，实线的偏移量和线宽；双虚线的（两）虚线间隔、起始位置、线宽、实部长和虚部长等。

实际制作线符号时，用户也要给出线符号绘制的基本参数，如定位线（即用户指定要绘制线符号的那条线或那些线）、线宽和线型等；GIS 系统根据用户要求，经过符号坐标和绘图坐标的变换，沿定位线不断地“铺设”，即重复配置组成单元。

重复配置点符号图元的线符号绘制方法比较复杂，因为它在将点符号单元沿定位线分段串接的同时，还需要将点符号单元在拐弯处作变形处理。串接或拐弯时处理的思路是：取点符号作为线符号的基本最小循环单元，求出点符号的外接矩形作为符号拼接或变形时参与运算的符号的有效范围；按符号长度在

定位线上分断截取，若外接矩形的长度超出拐点则截去超出部分，截去部分转到下一折线段内处理。

如果定位线是直线，绘制较为简单。若定位线是曲线，则算法比较复杂。

3. 矢量面符号绘制

面符号绘制基于上述两种符号绘制。一方面，多边形边界采用线符号绘制法；另一方面，采用点模式或线模式进行面域内部的充填。前者较为简单，因为多边形边界一般采用不复杂的普通实现或虚线。因此，面符号绘制的关键技术在于面域填充。

面域填充有三种情况：一是在绘图区域内以不同的倾角、不同的间距、不同的实、虚部长度的平行线族来构成不同的图案，即晕线填充图案。晕线填充涉及到倾角、线宽、起始位置、偏移量、实部长、虚部长和颜线色等参数。

二是在绘图区域内以不同的间距、不同的布点形式（井字形、品字形）和不同的旋转角绘制点状符号以构成图案，即点符号填充图案；它涉及到行偏移、列偏移、行间距、列间距、缩放系数、旋转角、点符号、旋转角形式（固定、随机）、存点形式等参数。

三是位图填充图案，它涉及的参数有位图长度、位图宽度、行间距、列间距、缩放系数、旋转角、填充形式（品字形、井字形）、位图等。

面域填充的关键是在面域内求晕线族，然后在晕线上绘指定线符号或按一定的距离绘制指定点符号。

求面域内晕线族可以采用这样的方法：首先写出确定水平晕线的程序。确定水平晕线就是按给定间隔给出水平线，再求解这些水平线与边界的交点。由于是水平线，算法比较简单。对于倾斜晕线，可以先将多边形旋转，使旋转后的晕线平行于水平坐标轴；采用上述水平晕线的程序求解晕线后，再将多边形进行同样角度的反旋转，即可得到倾斜晕线（图6-8）。

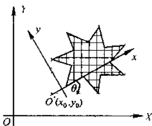


图6-8 面符号晕线

在GIS软件的地图符号库中，也有一系列面状符号，或填充图案，如图6-4所示。符号库中存贮的填充图案，其实也是组成它的点符号或线符号构成单元，以及其组合、配置的有关信息。实际制作面符号时，用户也要给出线符号绘制的基本参数，如指定要绘制符号的那个或那些多边形和指定的填充图案等。GIS系统根据用户要求，经过符号坐标和绘图坐标的变换，将点符号或线符号构成单元在多边形内重复配置，直至填充完整个面域。

6.2.3 地图符号库及其应用

前面业已多次提到地图符号库的。那么，什么是符号库，它是如何建立和组织的呢？

1. 地图符号库一般概念

地图符号库是多种地图符号的集合。符号库具有相关的管理功能，对不同类型的地图符号、注记类型、色彩方案和坐标系统等进行统一管理。商用GIS软件系统都带有地图符号库，除了库中包含很多基本的地图符号和相关要素外，还提供符号或符号库定制功能。

上述地图符号库是一个统称。在总概念下，不同层次的类别和不同比例尺的点、线、面和注记符号还形成众多具体的符号库。例如，图6-9是Arc/Info中符号库的一个检索状态，从图中左部的目录可见，此刻正检索到Arc/Info提供的地质专业的符号库，地质符号库下包含很多更具体的地图符号库。

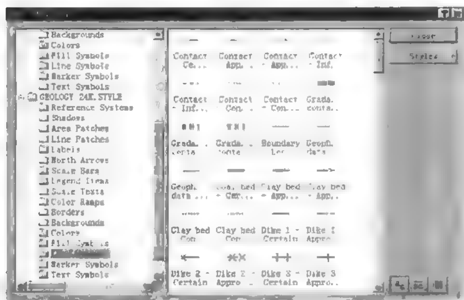


图6-9 Arc/Info符号库一瞥

符号库实际上是符号描述信息的集合。按现代GIS多采用的面向对象方法，无论是点符号、线符号还是面符号，都是各自图元对象类的实例对象聚集而成的复杂对象；组成符号的图元对象标识和图元描述参数或信息，作为一个个数据文件保存在符号库中。当采用间接信息法时，符号描述的参数或信息可以

通过交互式符号设计系统获得。

符号库不仅包含描述符号信息的数据文件，而且还必须建立索引机制。符号库信息的应用效率很大程度上取决于索引机制的好坏。

2. 地图符号库的检索机制

索引机制建立的方法有两种：一是索引与数据放在同一文件中，索引存放在数据文件之前，即符号库分索引区和数据区；二是索引与数据分开存放，即索引存放一个文件，数据存放在另一文件。前者的特点是一个符号库仅一个文件，便于管理，但不便添加符号。因此，这种方式一般用于设计系统的输出符号库组织。后者是一个符号库由两个文件组成，为便于管理，可采用文件名相同而后缀不同的方法组织符号库。

尽管点、线、面各种符号由于符号特征不同，索引结构会有所区别，但三种符号索引结构中都应包括：符号名、符号代码、描述数据指针、描述数据大小。

为了管理多比例尺系列、多种形式的符号库，不同的 GIS 系统有不同的实现方法。通常都将系统所涉及的符号库放在系统的一特定目录下。为方便对符号库的操作，有的系统指定当前操作对应的符号库，如 MGE 和 AutoCAD；有的系统则对存放在特定目录下的符号库建立索引表。例如，GeoStar 通过建立一个索引文件来实现符号库管理，该文件的一条记录对应一个符号库，用户每生成一个新符号库就在符号库索引文件中增加一条记录。符号库索引文件内容包括：符号库序号、符号库类型、符号库名称和符号库描述信息等。引入符号库序号的目的是为了方便记录符号所在的符号库，从而快速获得某个符号的描述信息。

当采用符号代码对符号进行索引时，为获得特定符号代码的符号描述信息必须顺序查找所有的符号库，显然要花较长的时间。如果采用符号所在的符号库和符号在符号库中的顺序号来提取符号描述信息，则要快得多。建立了这样的符号库索引文件之后，用户可以设计不同比例尺、不同类型的符号库，为 GIS 支持多比例尺制图输出提供了极大的方便。

符号库设计时，任何符号都应有一个符号代码，它是符号的唯一标识码。若 GIS 系统中地物编码和符号代码的关系要求一一对应，一般符号代码的设计以地物编码为准。相反，如果地物编码和符号代码的关系仅保持着一种松散的关系，即地物对应的符号可以随时改变，这时符号代码的设计可以按国家标准规定为依据，可以顺序编号，也可以按符号的大类、小类、识别码按一定的组合形成。

具体设计符号时，应注意考虑符号的精度、符号定位点、组成线符号的最小符号单元、侧向（不对称）线符号的侧向以及符号的颜色等。

3. 地图符号的二次开发

商用 GIS 软件系统本身带有基本的地图符号库，这些符号库面向 GIS 用户的一般需求，包含比较多的很多图例符号和相关要素，能满足用户的常规制图需要。但是，GIS 的应用领域很广，用户千差万别，特别是多样化的专题地图制作对符号会有很多特殊的需求，经常超出软件原有符号库的能力，不少符号，如拦江大坝、信号塔台、电力线等，难以实现。这时，用户就要应用 GIS 软件平台提供的定制地图符号或编辑符号库的功能，形成面向某种或某些专门应用的地图符号库。这就是地图符号的二次开发。

GIS 软件平台提供的定制符号功能，通常表现为一个符号编辑器，或符号定制系统，用户能够通过这样的系统方便地制作自己所需的地图符号。符号编辑器的要求界面友好、操作灵活方便，设计时能实时动态观察所设计的符号，符号设计能满足一定的精度，以及有良好的封装性、可维护性和可适应性等。

地图符号库通过符号编辑界面提供一些基本的地图符号或图形单元，它们是可以用来构成新的地图符号或图形符号的基本元素或单元。用户定制时，可用鼠标选择所需的基本符号或图形单元，基于这些图元，用户可以通过下述任一方式来定制新符号：或者在屏幕上基于这些图元来直接绘图，或者在图元参数输入对话框中输入有关参数，如线型、宽度和颜色等。在定制过程中，用户能观察到实时的变化或进展，随时进行修改或完善。用户还可对符号库中的符号进行删除、修改、浏览、选取等。为用户使用方便，还可以对符号库中的符号按符号的索引信息（如符号标识码）进行排序。

目前绝大多数国产和国外 GIS 软件都提供符号定制系统，有较友好的用户交互界面。新版的 Arc/Info 系统的符号设计系统可谓是一个很好的范例，使用非常方便，界面也很友好。图 6－10 是 Arc/Info 的 ArcMap 中提供的符号编辑界面（局部），用户可以很方便地调用系统内含的图元设计出自己的符号，甚至符号库。

4. 现有软件的地图符号系统的应用特点

目前国内外的 GIS 与 CAD 软件设计符号库的思路和实现符号定制的途径，可归纳为以下四种：

一是文本编辑器设计方法，如 AutoCAD 的形文件、线文件和阴影文件，设计速度慢，不能实时观察所设计的符号。

二是采用系统提供的二次开发语言编程实现，如 Arc/Info 的 AML 语言等，提供了编程实现符号绘制的接口，开发者可以利用它们实现特殊符号的定制。

三是利用系统本身的图形编辑功能实现。用户先在屏幕上绘制、编辑所要设计的符号，然后圈定符号范围并指定符号定位点，系统即从当前的编辑缓冲区中提取符号描述信息；当用户指定了诸如符号代码（或名称）等参数后，即可

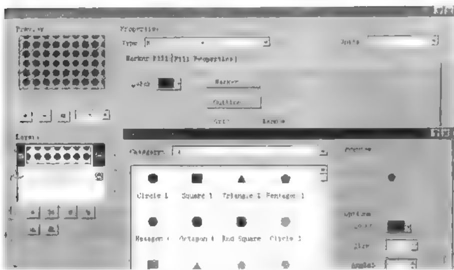


图6－10 ArcMap 符号编辑界面（局部）

存放到符号库中。如：AutoCAD 的块文件（block）和 MGE 的单元（cell）等，都是采用这种设计方式。但是，这种方式由于系统图形编辑功能的局限，只能设计点符号。

四是提供符号设计界面。例如，上述 ArcMap 中符号库编辑界面（图 6－10）能直接利用系统提供的各种符号，或拼装组合系统提供的基层简单图元，来完成针对各种专业需要的不同类型的符号设计，而且使用起来非常方便。MGE 提供的线型编辑功能也可用于交互式编辑线符号。

虽然不同 GIS 系统提供的二次开发环境有所区别，但目前绝大多数 GIS 系统提供了对数据库底层进行读写操作的 API 函数或控件，使用通用开发软件，如 VB、VC ++ 和 Delphi 等，开发人员可以很方便地集成这些函数和控件设计出符合特定要求的应用模块，如空间实体符号化模块。

AutoCAD 内含 AutoLisp 语言、ADS（AutoCAD development system）和 ARX 开发系统，便于用户进行二次开发。LISP 语言作为比较完整的编程语言，为用户提供了强大的二次开发工具，用户可以利用 AutoLisp 编制各种程序，从而为 AutoCAD增加新的命令，也可以为各专业编制图形库。ADS 是编写与 AutoCAD 一起工作的 C 程序的最初方法。ADS 程序在完成计算量较大的任务上，比 AutoLisp 程序要快，且提供了对操作系统的底层访问，但编写起来也更复杂。ARX 是编写与 AutoCAD 一起工作的 C 和 C ++ 程序的新方法。ARX 程序与 AutoCAD 的联系比 ADS 程序更加直接，它们对于诸如渲染和实体造型这样的非常

密集的任务能够以更快的速度计算。

为实现对系统的二次开发，Arc/Info 提供了宏语言 AML 和开放式开发环境 ODE。其中，用 ODE 开发基于 Arc/Info 的应用程序，可以使用标准的开发工具例如，在 Windows NT 或 2000 环境下，可用 VB，VC ++，Delphi 和 PowerBuilder 等进行开发，可以建立标准的功能强大的操作界面，支持所有的 Arc/Info 命令，执行效率更高。

模块化 GIS 环境 MGE 提供了 MDL 语言进行二次开发。MDL 采用标准语言格式，提供了 CAD 软件开发的集成环境、CAD 专用库、面向 CAD 的内部函数，使 MDL 成为强大的、丰富的、实用的开发系统。在 MEG 环境下，为获得一幅符合国家(地形图图式)标准的地形图，除设计相应图式的符号库外，还必须用 MDL 开发一些特殊功能的应用模块，包括特殊符号绘制子程序、设计 DGN 文件(MEG 的基本图文件，即工作区，见 §3.2.3.3)实体整体符号化程序、图幅整饰程序和注记压盖处理程序等。

§6.3 地图注记、排版和输出

地图注记也是地图符号系统的组成部分；地图的排版则是在地图符号化和注记之后，需在地图输出之前完成的工作。地图注记、排版和输出是三个不同阶段相互独立的内容，但由于本书的特点，几项内容不作详谈，因而合成一节一并叙述。

6.3.1 地图注记

地图注记包含两方面的内容，一是地图上的内容注记，主要是地名注记。注记借助语言文字的功能来实现，它本身可以作为空间数据库中一项内容。另一种注记是制图说明注记，这种注记仅与地图输出有关。空间数据库中一般不存储这些内容。

本节主要讨论前一种注记，制图说明注记放在下一分节(§6.3.2)的图幅整饰中讨论。

1. 注记的数据结构

在计算机内进行地图注记，可直接采用目前流行的印刷出版行业用的字库。这种印刷字库具有多种汉字字体，包括宋体、仿宋、细线、中等线、黑体、隶书、楷书等。每种字体约有 7 000 至 12 000 字左右，都有点阵字库、矢量字库和 TrueType 三类字体。点阵字符已从 24 点阵发展到 128 点阵，即每个字符包含有 128 × 128 个二进制位，点阵字符的汉字因此会占用大量存贮空间。矢量汉字外形光滑而且存贮空间较少，适合于地图使用。TrueType 字体是一种点阵与矢量

结合的字体,它以普通栅格字库的形式存放,但在显示时,用数学曲线拟合注记的边缘,外形美观。

一般用于地图输出的地图注记,数据结构都较复杂。不仅要考虑注记点的位置坐标,还要记录注记的字体、颜色、大小、注记间隔、注记的方位角和注记的字形(指左斜、右耸等),以及注记的内容等。一个好的注记系统应当既考虑到注记的组合方式,又考虑注记字符串中每个字在实际图幅中方向的变化。例如,为很好表现大兴安岭的注记整体,应沿着山脉注记,而且每个字的方向可能还要随具体山脉延展方向而调整。此外,注记的数据结构自然还不可缺少注记目标标识和所属地物类或层等因素,否则注记不可能标注到应相应的地物上。

综上所述,注记的数据结构应包含如下因素:注记目标标识、所属地物类或层、字体、颜色、大小、字形、字符间距、注记方式、字符个数、注记字符串、注记点位坐标 X、注记点位坐标 y 和注记点位方向角 A 等。可见,计算机中的地图注记包含丰富的内容。

2. 注记方式与注记编辑

注记的方式是注记中的一项重要内容。注记方式分四种:单点注记、双点注记、布点注记和参考线注记。单点注记是最常见的方式,该方式给定一个点的坐标,以此点为起点注记一串字符,注记方向角一般为零。双点注记则是指给定两个点的位置坐标,注记的一串字符及其方向由这两个点位来决定。第三种是布点注记,它将一个整体注记的每个字符对应一个点位和方向,其中单点布点时给定一个点位;双点布点时给定两个点位,方向由两个点来决定。参考线注记沿一条已有的参考线如公路线、平行或垂直于参考线自动布点,效果与双点布点相类似。

注记编辑经常性的内容是修改注记的字体、颜色和大小等选择参数,修改注记内容,移动注记点位,修改注记方式和注记方向等。此外,还要考虑注记相互压盖和注记与地物的压盖与协调问题(参见图 6-11)。

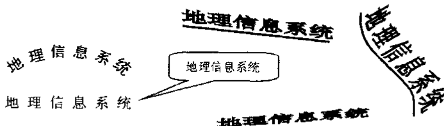


图 6-11 几种注记方式

3. 地名自动注记

地名是地图的最重要的组成部分之一，通过地名注记可以一目了然地识别出重要的地物目标，获取相关信息。众所周知，地图上的地名注记特别是小比例尺地形图等的地名注记，仅点状地名的注记一般就可达到几百上千个，工作量相当大。因此，地名注记的自动化一直是计算机地图制图研究寻求突破的一项重要内容。特别是从20世纪60年代起，GIS的迅速发展及其广泛的应用，对GIS图形输出提出了更新更高的要求，地名的自动注记问题因此而提到较重要的议事日程上来。

地名的自动注记非常困难的原因，关键在于正确处理地名注记与地理要素之间的关系，地名注记与其相应的地理要素关系密切，但没有固定的一一对应关系。地名注记在地图上的位置是不确定的，地名注记与其相应的地理要素只存在根据优先顺序和它与其周边要素之间的压盖情况进行配置的关系，它是由制图人员在绘制地图时根据制图规则配置，具有相当程度上的主观随意性，这相应地也增加了自动注记的难度。

目前，国内和国际上虽然对地名注记自动化进行了大量的研究，但是，距离建立一套能灵活运用于生产的、真正意义上的"地名注记自动化"实用系统，还有很大差距。至今为止，多数地名注记自动化的研究是基于"人工智能"算法，例如，基于"状态空间问题"求解的搜索策略和专家系统等。它们多是基于注记位置以及它与其他地理要素压盖情况制定的规则，进行搜索、推理。不论是"专家系统"还是哪一种算法，它们都是一种"串行算法"。问题的解与注记的顺序有关，后面的注记直接受到前面注记位置的影响。

武汉大学（原武汉测绘科技大学）的张祖勋教授等提出了一种基于Hopfield神经网络的整体最优解的地名自动注记算法。该算法的核心思想是：从整体（整幅地形图）上考虑地名注记与其相应的地理要素、各地名注记之间，以及地名注记与其周边要素之间的关系，利用Hopfield神经网络模型的能量收敛特性，获得一个整体最佳的注记结果——即整体最优解。

鉴于地名注记分为点状、线状和面状符号注记，三种注记各有特点，整体最优解算法对三种符号注记的数据处理给定不同规则。经过对点状、线状和面状符号注记的全面考虑，该算法选择点状符号"整体最优注记"为重点，选择神经网络中的Hopfield模型为整体最优配置的数学模型。线状要素通过提取要素平行线，面状要素通过提取要素骨架线加上一定约束条件后，转化为点状要素情形处理。

应用Hopfield神经网络模型首先要根据问题的物理模型建立一个能量函数。对于点状要素自动注记而言，所谓整体最优配置就是按点状符号配置的优先等级，并顾及所有的地图要素（包括注记本身）之间的关系，从每一个符号的

多个备选位置中选取一个最优解。其标志是能最佳地满足配置的规则和最少压盖。实验结果表明,用神经元网络整体最优解算法解决自动优化组合问题,包括注记自动配置最优解求取问题是可行的,许多性能优于"串行算法"。由于网络是以迅速收敛的迭代方式运行,所以,那些有"组合爆炸"危险的复杂问题可以变为可计算的简单问题,从而使算法效率有数量级的提高。

6.3.2 地图排版

地图排版指地图符号化与注记之后的颜色配置、图幅整饰、图例和图面布局等工作。

1. 颜色配置

一般的输出,如用绘图仪绘制纸质地图等,对地图的配色要求不一定很高。但是,地图如果要出版印刷,则色彩配置必须十分讲究。为了能自动化地提高色彩配置的效率和质量,必须首先找到地图色彩的数字表达方法,并建立地图色彩数据库,然后进行人机交互或自动设色。

颜色的配置需完成以下几项工作:

第一,定色度数据。定色度数据是指应用色差计测定专题地图色谱中的三套叠色系统和85个连续色表的色度数据。色度数据实现了地图色彩的数字表达和空间定位,为计算机处理色彩提供了定量依据。为了确保色度数据的正确性,一方面要注意对色差计进行标定工作,严格挑选颜色实物样本,并控制环境温度;另一方面又要根据色度学理论,设计专用于检验色度数据质量的图形处理软件,输出色度坐标图和色立体图,以发现和纠正错误的色度数据。

第二,建立地图色彩数据库。有了色度数据,为适应系统对地图色彩自动设计和数据查询的需要,还必须设计相应的地图色彩数据库结构并存储下列数据:

- 专题地图色谱的色块定位(查询索引)数据。即页码和编号。
- 印刷数据。指色谱中每个色块的网线合成比例。
- 色度数据。包括三刺激值、色度坐标、空间匀色坐标、主波长和光谱纯度。
- 分类命名编码。分类命名是依照地图设计人员的习惯,利用颜色的视觉心理效果或颜色的实物实样来挑选颜色。为此,需要对数以万计的颜色依照其视觉心理效果加以分类和命名编码。
- 视觉心理效果的对比、感觉和象征编码。指颜色的对比(互补、对比和类似)、颜色的感觉(如冷暖、干湿、喜忧等)和颜色的象征(如喜庆、严肃、灾害、病态等)。

上述第四点中的分类编码应与国际孟塞尔分类系统方案相衔接。例如,适

用于黄、品红、青叠色系统的我国四级分类命名系统是：

- 色系。红、橙、棕、黄、绿、蓝、紫。
- 色相。品红、紫红、红、红橙、橙、黄橙、橙黄、黄、绿黄、黄绿、绿、蓝绿、绿蓝、青、蓝、紫蓝、蓝紫、紫、红紫。
- 根据颜色的鲜艳与灰暗度，将每个色相区的全部颜色划分为：鲜艳区、次鲜艳区、次暗区、暗区和非彩色区（黑、灰、白）。
- 根据颜色的深浅度，将每个色相中的全部颜色划分为深色区、中色区和浅色区。

第二，需要建立荧屏色与印刷色之间的对应关系

GIS 制图中，时常发生荧屏上制作满意的地图作品，在硬拷贝输出时色彩却不如人意的情况，这是由于荧屏色和印刷色的产生原理和载体有所不同，荧屏上的颜色效果和印刷色有一定出入的缘故。印刷色由黄、品红、青、黑四种原色油墨，按减色法原理合成；而荧屏色由红、绿、蓝三种荧光粉原色，按加色法原理合成。再加上在不同型号的荧屏中，产生颜色的电子元件具有不同的技术指标，如灰度等级、分辨率等的差异很大，致使荧屏色模拟印刷色带来许多困难。实验结果表明，印刷色与荧屏色中的某些颜色存在一定的差别。例如，荧屏色的色域范围不能覆盖印刷色中的黄和青色相中的某些颜色，而荧屏色中的 R、G、B 原色中的某些颜色却分布在印刷色的色域范围之外。荧屏色与油墨色两种颜色系统之间的差别大致如图 6－12 所示。

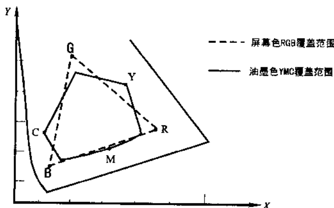


图 6－12 荧屏色与油墨色覆盖范围比较

为了有利于得到满意的硬拷贝输出地图，必须建立荧屏色与印刷色之间的对应关系。为此，可以利用目视和仪器监测相结合的方法：首先在荧屏上产生黄、品红、青三种印刷油墨原色的对应荧屏色，记下它们的 R、G、B 值。然后，使

用颜色数据库中的三刺激值和色度坐标，迭代计算出色差最小的转换系数矩阵。根据这个关系，可产生荧屏色谱和荧屏色 RGB 值数据库，并使它们与印刷色谱和印刷色数据库建立一一对应关系。基于这种对应关系，即可通过荧屏色实现"所见即所得"的目标。

2. 图幅整饰

一幅完整的地图除了包含反映地理要素的线划及色彩要素以外，还必须包含与地理数据相关的一系列辅助要素，如图名、图号、图表拼接、方里网、图例、比例尺、坡度尺、图幅外注记等诸多内容。图幅整饰就是完成以上这些辅助要素的操作。普通地图的图幅整饰内容和注记大小等，均有严格的规定，必须符合规范和图式要求。但对于专题图而言，图幅整饰就随意得多，它以合理和美观为原则。

专业 GIS 软件一般都为用户提供方便的图幅整饰工具。图 6－13 是 Arc/Info 为用户提供的制作图例(legend)的人机交互界面，它是 Arc/Info 的图幅整饰工具的一部分。图中可见，界面上部有 4 项菜单，除此刻(图 6－13)打开的 legend 项外，还有三项：items 指图例的符号，打开 items 项有不少制作图例符号的工具；frame 指图例制作的外图框；size and position 则是管理图例大小和位置的功能。这些工具非常方便于用户进行图幅整饰。

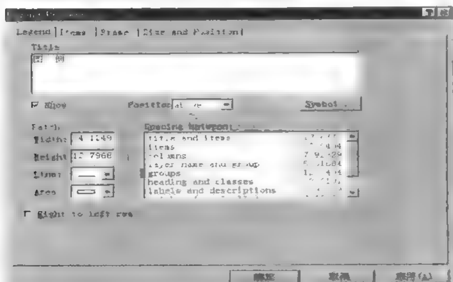


图 6－13　Arc/Info 的图幅整饰工具

图 6－14 是一个经整饰、待输出的专题地图实例。

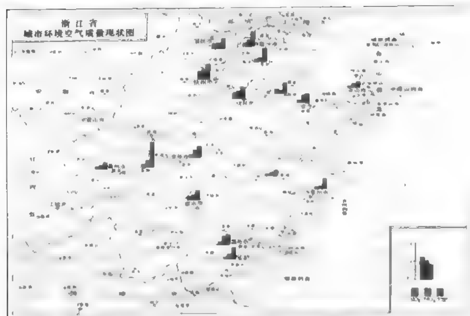


图 6－14　专题地图示例

3．地图排版布局

在地图输出到介质之前最后一步要做的就是地图排版布局。之所以需要进行这一步工作，是因为地图常常需要与描述它的其他附加性材料，如各种附图、统计图表、图片和文本等，一块儿排版输出。例如，我们在随处可见的各地交通旅游地图上一般都可以看到一系列关于饭店、旅游景点和各种公共交通信息的图、表和文字等。因此，在地图输出前，需要将地图与附加性图、表和文字等在计算机中混合排版，形成统一的绘图文件，为是除纸质地图、电子地图和分色加网软片做好准备。这里，排版布局的好坏很大程度地影响着地图的美观和可用性。

现代 GIS 软件都有将地图与附加性图、表和文字等集成在一起的能力，并能形成各种绘图仪和激光照排机所能接受的文件。几年前，GIS 软件在这方面还有不少困难。近年来得益于计算机技术的飞速进步，GIS 软件的输出排版能力进步很快。特别是微软的对象链接与嵌入技术，能把不同的文件都作为一个 OLE 对象，将它们集成在一起并按要求排版布局，最后建立统一的输出文件。

6.3.3 地图输出

1. 地图输出系统

至此,我们讲述了地图最后输出前的绘制全过程,包括建立地图数学基础、地图的符号化、地图色彩编辑、文字编辑(注记)、地图整饰和地图排版布局。现在来介绍除荧屏输出以外的地图输出。

地图输出的途径主要有三种,也可以说有三种主要的地图输出系统。一种是绘图仪或打印机输出,地图输出到硬拷贝(纸张、薄膜等)上。第二种是自动制版系统,这时地图输出形成分色加网软片,以供印刷之用。第三种途径是电子地图(集)系统,输出的是电子地图产品或电子地图集。下面分述这三种输出方式。

2. 绘图仪或打印机输出

绘图仪输出是最简单的,也是最常用的地图输出方式。目前最广泛应用的绘图仪,是彩色喷墨绘图仪,该仪器的性能价格比在近十年来有明显提高。近年来,小型的 A4 幅面彩色喷墨打印机的性能价格比,也提高很快,目前价格已降至千元以下,而输出质量对一般民用已而言已足够精美,因而被广泛用来作为 16 开本大小级别的文件、报告等的彩色地图插页输出。从原理上看,彩色喷墨绘图仪和彩色喷墨打印机的绘图机制相近。它们不仅性能价格比高,而且应用非常方便,因为过去 GIS 软件公司要针对不同的绘图仪编写不同的绘图驱动软件,而现在这一工作逐渐标准化,这些工作均由操作系统提供的驱动软件,或绘图仪、打印机生产公司提供的驱动软件完成。

从软件指令上看,绘图仪输出的实现方式有三种方式,一种是根据绘图指令,编写绘图程序,直接驱动绘图笔绘图。第二种方式是由 GIS 软件产生一种标准的图形文件,如 Windows 的元文件 WMF 文件,调用操作系统或者 Windows 提供的函数"播放"元文件,来绘制地图。第三种方式更为简单,所有程序不变,仅在需要绘图时,将图形屏幕显示的句柄改为绘图设备句柄即可。第一种方式即为笔式绘图仪输出方式,这种方式现已少用(参见 § 2.2.4.5)。彩色喷墨绘图仪和彩色喷墨打印机采用第三种方式。

3. 自动制版输出

自动制版输出包括分色加网处理技术和栅格影像处理两种方式。

分色加网处理主要分分色处理和加网处理两步骤,前者将已获得的彩色地图文件按照每一种颜色的黄、品红、青、黑的实际构成比例进行分色;后者则根据印刷彩色地图的网目密度进行加网处理,为输出分色加网胶片完成预处理工作,即产生一种包括符号和正文处理的国际上通用的标准格式文件。这种单色文件可以通过影像曝光机输出加网胶片。

为了完成分色处理，通常建立色表，这是一个工文文件，建立绘图文件中的笔号和不同的页面描述方法之间的联系。加网过程实际是一种填充方法，由于要求的网目密度通常在150以上，因此其加网过程需占用很多时间，应当寻求一种快速算法，postscript语言提供了较好的填充算法。

常用以建立色表的模型有：RGB三原色模型、HLS、色度、饱和度、亮度）模型和CMYBk（印刷油墨色的黄、品红、青、黑）模型。R、G、B值的变化范围是0至255；H、L、S的取值范围从0至100，印刷油墨色使用网目构成百分比进行分色处理。此三种模型的色表格式大体相同。

栅格影像处理将转换矢量式的页面描述文件（postscript）为点阵式影像文件，它可直接用于输出网目片或正文、符号和线画软片，从而完成印前处理的最后一步工作。转变过程中，需要计算网目尺寸和扫描线的匹配关系。RIP软件直接接受postscript文件并进行解释和转换工作，转换后的结果通常可适用多种型号的影像曝光设备。RIP软件直接接受矢量式文件，因而可以获得光滑的点阵边界。由此缘故，过去采用直接点阵式数据输出的方式正逐渐被淘汰，被postscript – RIP方式所替代。RIP过程中可以设置页面大小、网目形状、网目密度、正负网点选择等。

4. 电子地图制作

电子地图是一种数字化了的地图，电子地图一般存放在磁带、光盘等的数字存储介质上。电子地图一般操作界面比较方便，不同的电子地图具有相对统一的界面，而且电子地图大多连接着属性数据库，能作查询、计算、和统计分析。

电子地图通常是系列化的，表现为电子地图集产品。电子地图集作为传统地图集的补充，它不仅存贮了地图集的全部内容，而且形成了一个信息系统，称为电子地图集系统。图6 – 15为杭州市的一个电子地图集系统。图中右上侧有一系列菜单，可供用户进行各种主题，如购物、交通、娱乐和旅游等主题的选择和切换。右侧中间栏给出主题下的各种专题，如交通主题下有火车、汽车和室内公交等的选择。右侧下方为相应专题中各要素的列表。而主视窗的下方为各种工具栏按钮，如要素查询、丈量、最佳路线显示和地图漫游等，可以通过它来选择相应的操作，实现相应的信息查询等功能。

从图6 – 15中不难理解，现代电子地图集系统已经远远地超出了传统地图集的功能。电子地图集系统丰富的功能可归纳为以下几点。

• 地图构建功能。由于电子地图有很好的交互性，不仅允许用户根据自己的需要和设计方案选择或调整地图显示范围、比例尺、颜色、图例和图式等，而且提供了更新或再版地图集的新方法。系统提供了强有力的地图和数据输入、编辑和输出改变的能力，能确保及时地更新数据，保持电子地图集的现势性，并为再版地图集创造了十分优越的制图环境。

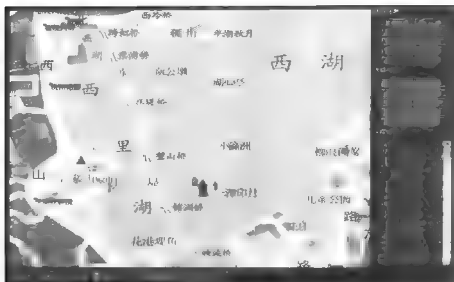


图 6-15 一种电子地图集系统界面示例

- 读图功能。读图功能包括点题、显示、闪烁、变色、开窗、对比等。
- 检索和查询功能。电子地图集系统可以根据用户需求来检索图中数据和属性，并以多媒体、图形、表格和文字报告形式提供查询结果。
- 统计和分析功能。电子地图集系统可以实现快速汇总，并进行统计分析，打印直方图，提供分组分类间隔选择和精度评价。
- 专题覆盖功能。电子地图集系统还能提供多种分析、覆盖方法，用以编制评价图、区划图、合成图和预测图等综合性地图。
- 制图综合功能。电子地图集系统中按视觉限度原理体现制图综合性能。这也是我们在§4.1.3.1中提到的屏幕上缩放地图时保持适当的图面负载量的功能。事实上，当数据库中存储了十分详细的制图数据时，正常位置的屏幕不可能显示全部图形细部。这样，当显示比例尺缩小时，必须抑制更多的细节，面状符号可能显示成点状符号；只有当使用开窗技术局部放大显示时，才有可能逐步显示出全部细节；多次放大就可获得多种比例尺的读图效果。为了实现这种制图综合的效果，细致地设计地图符号及其颜色是十分重要的。

由于电子地图集综合了一般地图集的功能和计算机图形系统二者的特点，电子地图具有许多的优点，包括交互性、现势性、灵活性、精确性和共享性，还能方便地实现无级放大、图幅的无缝接合，信息量丰富等等。

电子地图的制作比较灵活，既可以采用专门的电子地图制作软件，也可

采用现有的GIS软件来生成电子地图的画面文件，然后用适当的软件，将这些画面文件集成起来，形成电子地图集。

第七章 GIS 新技术发展

地理信息系统伴随着计算机技术的发展走过了几十年的历程，经历了 个发展阶段，产生了三代 GIS 软件，由最初的单机绘图、简单属性数据管理等功能变成了图形图像处理功能强大、大型数据库支撑的集成与分布相结合的信息系统平台。近几年来，GIS 研究与开发又取得较大进展，随之产生了一些 GIS 新技术分支。其中具有代表性的技术有：互联网 GIS（万维网 GIS，Internet GIS、Web GIS）、虚拟现实 GIS（VR GIS，reality virtual GIS）、组件 GIS（COM GIS）、时态 GIS（Temporal GIS）、空间数据库引擎（spatial database engine）技术等。本章将分别讲解这些新技术。

应当说明，本章内容并非地理信息科学及技术新发展内容的全部，之所以讲解主要的新技术，是因为作者在实践中感到有较多的 GIS 用户对这些新技术感兴趣的缘故。他们很想了解近年来出现的这些新技术及其新名词是怎么回事，希望能有一个较详细的全面的介绍。全章内容都是介绍性的，均不作基本要求。

§7.1 万维网地理信息系统

互联网（Internet）自 20 世纪后期出现以来飞速发展，现已几乎进入了人类社会生活的各个领域，对社会文明进步和经济发展产生了极为深远的影响，Internet 技术正在改变着整个世界。进入 21 世纪后，信息技术发展更加迅猛，随着通信、视频和宽带等信息网络与 Internet 相互融合步伐的加快，以及下一代互联网 Internet 2 技术的日趋成熟，一些影响互联网普及和进一步应用的技术制约因素将得到解决，互联网日益成为信息化社会信息交流、信息获取的重要工具。基于 Internet 的 browser/server 体系结构的应用模式已经成为一种新的工业标准，被广泛用于信息的发布、检索等诸多领域。毫无疑问，互联网已经成为当今世界最大的信息网络。

GIS 技术的飞速发展虽然为地理信息的电子化、可视化、网络化和中央存贮管理化带来了重大革新，但地理信息只限于局域网络内部使用。而社会对地理

信息的需求在不断增长。Internet技术的迅速发展为GIS提供了一种崭新而又非常有效的地理信息载体，尤其是“数字地球”概念的提出，引起了人们对GIS技术的广泛关注，已经成为新的研究热点。这就使得Internet环境下的空间信息处理技术成为实现“数字地球”的关键支撑技术之一，因而受到整个信息领域的高度重视。因此，GIS的网络化应用趋势已成为必然，以单机或局域网络为操作平台的工作模式已不能满足用户的需要。

7.1.1 Web GIS概念

1. Web GIS基本概念

万维网地理信息系统(Web GIS)指基于Internet平台、客户端应用软件采用WWW协议，运行在万维网上的地理信息系统。人们也时常将这一技术通称为互联网地理信息系统，即：互联网GIS，因特网GIS和Internet GIS。它是利用互联网技术来扩展和完善地理信息系统的一项新技术，其核心是在地理信息系统中嵌入HTTP和TCP/IP标准的应用体系，实现互联网环境下的空间信息管理等地理信息系统功能。由于Web仅是基于互联网技术而发展起来的一种技术，从更广泛的意义上讲，Web GIS仅是互联网地理信息系统(Internet GIS)中的一种。随着技术的进步，客户端也可能采用新的应用协议，产生新协议环境下的互联网地理信息系统。总之，互联网GIS是地理信息系统技术和互联网技术相结合产生的一种崭新的革命性的新技术，使基于地图(图形、图像)的应用系统得以通过互联网技术在各行各业中得到广泛的应用。

Web GIS是由多主机、多数据库与多台终端，并通过Internet/Intranet连接而组成的。实际上Web GIS常常是通过Internet连接大量的，分布在不同地点的不同部门的独立的GIS系统组成。由于Web GIS具有客户/服务器结构，因而客户端具有获得各种空间信息和应用的功能，在服务器端系统向客户端提供信息或系统服务。对于GIS应用来说，人们要求在互联网上不仅能够浏览文字信息，而且需要浏览大量的动态交互图形信息，并且不同的用户有着不同的需求。静态的主页不能满足要求，需要采取双向交互式方法的动态主页来实现，即根据用户的指定从数据库里查寻数据生成主页，分发给用户。

2. Web GIS组成

Web GIS在结构上属于分布式地理信息系统模型，通过Internet / WWW机制可有效地实现分布式地理信息处理。Web GIS一般由四部分组成：

- Web GIS浏览器(browser)。它可以通过Web服务器连通到任何地点的另一个数据服务器上，读取各种地理信息；
- Web GIS信息代理(information agent)。空间信息网络化的关键部门，主体(agent)是信息代理机制和信息代理协议，提供直接访问数据库的功能；

• Web GIS 服务器。它能解释中间代理请求及操作数据库服务器和实现浏览器和服务器的动态交互；

• Web GIS 编辑器(editor)。具有可视化、交互式、多窗口的功能，能建立 GIS 对象模型和进行空间数据的编辑及显示。

3．Web GIS 的应用及意义

Web GIS 开拓了地理信息资源利用的新领域，为 GIS 信息的高度社会化共享提供了可能，为 GIS 信息的提供者和使用者提供了有效途径，为传统地理信息系统的发展提供了新的机遇。它改变 GIS 数据信息的获取、传输、发布、共享、应用、可视化等过程和方式。互联网为 GIS 数据提供者在 WWW(world wide web)上提供了方便的 GIS 数据信息发布与共享方式，大量的地理信息系统的信息在互联网(Internet)上以 WWW 形式发布。这样的信息正以在惊人的速度增长。很多 GIS 数据信息可以在 WWW 上获得。在线的空间数据仓库、库目录为 WWW 用户提供在线数据服务，使任何地方的任何 Internet 用户都可获得他们所感兴趣的地理信息，使地理信息真正成为整个社会的共同财富。用户不必买昂贵的 GIS 软件，可直接通过 Internet 获取 GIS 数据和使用 GIS 分析功能，以满足不同层次用户对 GIS 数据的不同需求，例如原始数据、中间结果或最终结果等。万维网与地理信息系统的结合 Web GIS 是 GIS 软件发展的必然趋势。事实上，万维网已经成为地理信息系统新的操作平台

7.1.2 Web GIS 主要特点

1．基于 Internet/Intranet 标准

Web GIS 支持 Internet 网络通信和 TCP/IP 和 HTTP(超文本传输协议)，采用标准的 HTML 浏览器作为应用外壳。支持通信标准对 Web GIS 来说是至关重要的：支持 TCP/IP 和 HTTP，就是意味着 Web GIS 能与任何地方的数据相连，不论是单位内部或外部。实现这一层次的网络协议标准化是实现其他所有功能需求的基础和前提，也是 Web GIS 结构优越性的前提。Web GIS 体系结构支持的各种协议标准如表 7-1 所示。

2．分布式服务体系结构

分布式服务体系结构是在客户端和服务器端都能提供活跃的、可执行进程的体系结构；它能有效地平衡两者之间的处理负载。诸如动态提取数据子集并进行分析的进程任务，一般应当在服务器端，而不是在客户端执行。空间信息查询集的选定和按比例缩放地图的工作则适合在客户端执行。这种在客户机与服务器之间的进程分布式处理，最大限度地发挥了现有计算机硬件资源的利用率。把数据量集中的任务放在服务器上，使得应用程序能支持其他的网络请求。分布式处理显著地降低了带宽要求并提高了系统的性能。它允许用户嵌

表 7-1　Web GIS 体系结构支持的各种标准

基础技术	
网络通讯协议	TCP/IP
文档和文件传输	HTTP
文档显示与应用程序集成	HTML
应用程序构造	
客户端集成	Plug-in、ActiveX、Java Applet
服务器端集成	CGI、服务器 API、Java
应用程序扩展	
客户端扩展	HTML、JavaScript、VBScript
服务器端扩展	CGI、服务器 API、Java

入自己定制的 GIS 服务，使用的数据既可以是本地的也可以是分布的数据集，从而使传统 GIS 向分布式 GIS 转变。

3. 高效利用空间数据资源

Web GIS 能充分利用已有的 GIS 数据资源和属性数据库数据，将常用的多种 GIS 数据转换成自己的空间数据格式和相应的关系数据库，保护用户的先期投资。服务器端的 GIS 数据（包括图形和属性数据）不需要全部集中在一台机器上，可以分散安装在不同的多台机器上，这些机器可分布在空间距离很远的地方，只要通过 Internet/Intranet 相联就可以。这种分散存贮数据的方式对于降低系统负载，加快访问速度，降低成本等方面都是很有帮助的。这是 Web GIS 一个极其重要的优势。

遍布全球的代理商可以直接为用户发布数据并提供服务。用户可以将广泛分布的数据和本地数据结合在一起，使不同地区的计算机主机协同工作。这种技术使得全球的存储在 GIS 数据库中的现有的空间信息发挥出巨大的效力。ESRI 的用户将可以在任何时间，任何地点共享和使用彼此的数据。任何人通过一个简单的浏览器界面就可以访问经过复杂的专业 GIS 分析产生的数据。通过不断提高的访问 GIS 信息的能力，各种组织和全社会作为一个整体将更有效地利用现有的空间数据资源。

4. 其他优点

Web GIS 的发布速度快，范围广，维护方便。由于运用了 Internet 技术，Web GIS 的信息更新之及时，发布速度之快，发布范围之广，是其他传统地理信息系统难以比拟的。Web GIS 的体系结构包括许多应用服务，如制图、查询、地理编码等。传统的地理信息系统在用户规模有所扩大，数据有所变更之后，都需对

原有系统做大量的改动。而在 Web GIS 中,则只需维护服务器端的一套数据,用户端都能及时看到更新的数据。

Web GIS 还具有友好的用户界面。它使用标准的 Internet 浏览器作为用户使用界面和工具,通过与用户交互可定制网页。开发工具丰富,功能强大,所开发的用户界面具有较强的多媒体效果,甚至使人获得虚拟现实的感受。并且操作简单明了,形象直观,使得一般用户也能使用。

Web GIS 系统的建设投资少。利用 Internet 的基础设施,以较少的投资就可以建立一套覆盖整个企业或全行业,甚至世界范围内的空间信息发布体系。终端用户不需要购买任何专门的 GIS 软件,就可以享受到真正的、实时的 GIS 信息服务。

在系统安全性方面,有的 Web GIS 软件(如 AutoGuide)还具有对数据访问的安全控制。通过口令密码可以限制访问人员的范围以及可访问的内容。对于面向全社会的专业信息系统,在实际应用中,往往要求根据不同的用户(如行政首长、各级部门工作人员、一般用户等)提供相应的信息。

7.1.3 Web GIS 基本原理和构建方法

1. Web GIS 基本原理

从网络技术角度看,Internet/Intranet 可以看作是以 TCP/IP 为通信协议标准、DNS 域名服务和 SMTP 简单邮件传输协议为基础、WWW 和 FTP 服务为支撑、实现多服务器和多平台的相互连接的计算机通信网络。目前互联网已成为企业或部门内外各种信息管理和交换服务的主要平台。从系统运行和技术上分析,GIS 系统要成为网络化的分布式处理的信息系统,必须符合或兼容 Internet/Intranet 相关的技术标准,包括支持 TCP/IP 网络通信协议、HTTP 文档与文件传输协议、文档显示与应用程序集成(HTML 和 HTML Browser)、服务器端集成(Web Server CGI 或 API)、客户端扩展(HTML,Java 和 Java Script)、服务器端扩展(CGI 和 Server API)、可执行的客户端与服务器后台进程的逻辑分布、多线程可扩展服务器、支持动态访问、强大的客户部件、远程制作(包括可视化远程制作工具)和图形文档支持嵌入的 URLs 等。

支持 Internet/Intranet 通信技术标准,对于 Web GIS 来说就是实现客户端与服务器端数据传输通信。Web GIS 系统充分利用了分布式体系结构 Client/Server 的技术特点,是 Client/Server 结构特点的一个典型应用实例。分布式应用体系结构能实现在客户端与服务器端都具备提供功能强大、可执行进程的体系结构,达到真正有效地平衡客户端与服务器端之间的处理负荷,实现计算分布和数据分布的目标,使系统具有可互操作性。从而可以把数据量集中的处理任务由服务器端执行,在客户端则完成诸如空间查询、专题地图生成等进程,以

充分发挥客户机与服务器各自的优势，最大限度地发挥应用系统的作用。Web GIS 系统的客户端/服务器体系结构使得系统具有良好的开放性，它能使系统软硬件资源共享、数据多重应用、跨平台运行、易于集成等特性。

可以说，Web GIS 系统是一个建立于 Internet/Intranet 之上的，采用开放式结构、具有统一标准和广泛适应性的网络应用系统。Web GIS 系统的客户端用户不仅能使用服务器端的硬软件资源，而且可以方便地利用 Internet 主干网的数据信息资源等。

一般将 Web GIS 分为两类，根据主要的图形属性数据所处的逻辑位置不同，可划分为服务器端和客户端两种解决方案，也有人将其分为动态式和主动式 Web GIS。有些 Web GIS 平台采用介于其中的平衡方案。对于服务器端的 Web GIS 解决方案而言，空间分析和输出全由服务器完成。它能解决大型数据库的管理及完成复杂任务，客户端仅负责用户请求和数据输入，通用网关接口 CGI 负责实现与 GIS 服务器的连接，最后客户端将结果显示出来。这种模式对网络传输无疑是非常繁重的负担。

对于客户端解决方案，一部分常见的 GIS 分析和数据处理工作在客户端完成，而不像服务器端解决方案，几乎包办处理用户的一切请求。系统需通过服务器向客户端发送一段运行在本地机上的客户程序。这个程序可以与用户相交互，处理用户的一些简单请求，如地图的开窗、放大等，所需的矢量地形数据直接向服务器申请。当客户发出一些较复杂、高级的操作要求而客户程序不能处理时，才请求 Web GIS 服务器处理，其处理结果也以矢量数据的形式发回给客户端。与服务器端解决方案相比，客户端方式的 Web GIS 具有用户操作灵活方便的特征，有效地减少网络传输和服务器的负担，但处理大型数据库和完成复杂的 GIS 空间操作的能力十分有限。

Web GIS 系统开发最初采用通用网关接口技术，以及相继发展起来的 IDC、ASP、ISAPI、NSAPI 等技术方法；后来又产生了被称为“插件”(plug－in)的应用技术，以及现在比较流行的 Java 语言，包括 Java 小程序 Applet 和 Java 虚拟机技术(VRM)等，和采用 ActiveX 控件及 COM 组件对象模型等网络开发技术。目前采用这五种技术方法构建的 Web GIS 系统平台在许多领域都得到了很好的应用。下面分述之。

2. 利用 CGI 技术方法构建 Web GIS 系统

通用网关接口 CGI(common gateway interface)方法就是互联网络服务器(Web Server)通过调用外部应用程序的接口扩展网络服务器的功能。这时 CGI 的作用即相当于在外部应用程序与 Internet/Intranet 网络服务器之间架设一座桥梁，使网络服务器对客户端的请求作出响应。客户端通过网络服务器激发 CGI 程序的响应实现具体的操作，读取超文本标识语言(HTML)文件，并将读取

的数据信息或文件，通过服务器送往客户端。

通用网关接口 CGI 的实质性作用是定义服务器和网关程序如何通信。在客户端，Web 浏览器以 HTML 建立用户界面；在服务器端，GIS 软件通过 CGI 与 Web 服务器相连。当用户发送一个请求到服务器上，服务器通过 CGI 把该请求转发给后端运行的 GIS 应用程序，由应用程序生成结果交还到服务器上。服务器再将结果传递到用户端。其工作方式如图 7-1 所示。

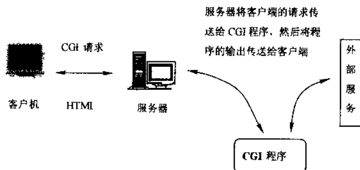


图 7-1 CGI 工作原理

根据这一原理，从理论上来讲，任何一个地理信息系统软件都可以通过 CGI 与万维网连接起来，远程用户可以通过浏览器发出要求，服务器把这一要求传递到后端的 GIS 软件中，GIS 软件按照要求产生一幅数字图像（GIF 图像），然后传回给远程的用户。

目前市场上推出的 Web GIS 系统软件，有一部分就是利用这一原理实现的，如美国 ESRI 公司的 Internet Map Server for ArcView 和 MapInfo 公司的 MapInfo MapXtreme 等。这种实现方式要在目前的 Internet/Intranet 浏览器上发布和传输 GIS 数据信息，需要解决两个技术难点。一方面是现有的网络浏览器不能读取矢量图形数据，矢量数据在网上传输需要先在服务器端转换成栅格图形数据，如 BMP、JPEG 等，因为现有 Web 浏览器仅支持栅格图形。这样的转换使得图形数据量成倍增大，并使本已拥挤的网络不堪重负；另一方面，传统的 GIS 原有的数据类型与 Internet/Intranet 现有的数据类型相距甚远，尤其是矢量图形与其属性数据已建立的对应、关联关系本已十分复杂，要在浏览器上实现原有的许多操作变得很困难。但采用这种技术方法构造 Web GIS 具有简单易行的特点，而且目前网络技术已有了较大的改善。这种方法通过采用原有 GIS 图形内核较易实现，适用于对原有 GIS 系统的网络化改造。

3. 利用服务器端应用程序接口建立 Web GIS 系统

服务器应用程序接口技术是为克服 CGI 方法的低效率问题而研制的，这

种方法的基本原理与CGI类似,所不同的是CGI程序是可以单独运行的程序,而基于服务器应用程序接口的程序必须在特定的服务器上运行。它的特点是速度要比CGI方法快得多,因为基于服务器API的动态连接模块启动后会一直处于运行状态,而不像CGI每次都要重新启动。

这种方法的缺陷在于它依附与特定的服务器和计算机平台。例如,微软公司的ISAPI只能在视窗(Windows)平台上运行。所以尽管CGI技术似乎显得有些落后,但由于其普遍性,在任何平台上都可以应用,而且可以用任何语言编写,其应用仍然非常广泛。许多商业化系统如ESRI的Internet Map Server(IMS)仍采用CGI技术。

总起来说,这里的服务器应用程序接口方法与上述CGI方法,都增强了用户端交互性,使用户可以获取各种地理空间数据和地图,但传给用户的信息仍然是静态的。用户不能操作单个地理实体以及快速放大和缩小地图,因为在用户端整个地图图像是一个实体,任何用户的GIS操作,如放大、缩小等都需要服务器来完成。当互联网流量较高时,系统反应会很慢。解决这一问题的方法之一是把一部分服务器的功能移到用户端。这样不仅可以大大加快用户操作的反应速度,而且也减少了互联网上流量和服务器的负载。

4. 利用plug－in插件技术方法建立Web GIS系统

插件法(plug－in)就是一种偏重用户端的、由美国网景公司(Netscape)开发的增加网络浏览器功能的方法。这种方法构造Web GIS系统的思路和原理与CGI技术方法有许多相似之处,所不同的是plug－in技术方法是在客户端的浏览器上增加一个能识别矢量图形数据的插件。通过这样的插件,使得服务器端的矢量图形数据无须转换,就能直接通过Web浏览器实现图形浏览、查询、分析等操作功能,大大减少了网络的数据传输量,较好地解决了网络上图形数据信息的传输"瓶颈"。同时,矢量图形与其属性数据已建立的对应、关联关系也易得到保存。

plug－in提供了一套应用程序接口(API),可用于研制与网络浏览器直接交换信息的专门的Web GIS软件包。插件使Web页面在现有技术标准环境下,可以随时加进去新的内容。音频、视频、图形、多媒体和商业与客户应用程序等全部是HTML页上的有效内容。目前流行的互联网浏览器,如网景公司的Netscape和微软公司的Internet Explorer(IE)等均具有应用程序接口(API),其目的就是方便网络开发商和用户扩展的与网络相关的特定应用。

世界著名的CAD软件厂商——美国Autodesk公司的MapGuide就是基于这一原理的Web GIS系统平台。这一系统利用位于客户端的MapGuide plug－in插件(插入件为Viewer)和服务器端的MapGuide Server,通过其特有的"地图窗口文件"(MWF)实现基于矢量图形数据的各种操作和管理,如动态发布与图

层管理等。地图窗口文件 MWF 包含地图属性、安全信息、地图图层属性、原始地图数据和用户接口标准等信息。这种技术方法的特点是以通用的浏览器为载体或平台，易于操作使用。同时它是嵌入式的插件，它自身所提供的强大的图形及数据库操作功能与浏览器功能相结合，较好地解决了图形与属性数据的双向浏览、检索、查询，以及统计分析等操作功能。美国 Intergraph 公司的 Web Map Server for GeoMedia（插入件为 ActiveCGM）也是采用 plug－in 插件技术方法实现的 Web GIS 系统软件。

插件方式建立的 Web GIS 系统的主要特点是速度快，客户端操作的是矢量地图数据。但它和传统的应用系统相比，需要先从互联网上获取特定的插件，安装后再使用。

5. 利用 ActiveX 控件和 DCOM 组件对象模型技术建立 Web GIS 系统

微软公司的 ActiveX 是一种对象链接与嵌入技术（OLE），可应用于 Internet 的开发。它的基础是 DCOM（distributed common object model）分布式组件对象模型。DCOM 本身并不是一种计算机编程语言，而是一种技术标准。组件对象模型 DCOM 和 ActiveX 控件技术方法具备构造各种 GIS 系统功能模块的能力。利用这些技术方法和与之相应的 OLE（对象链接与嵌入）、SDE（空间数据引擎）技术方法相结合，可以开发出功能强大的 Web GIS 系统。

美国 ESRI 公司于 1997 年推出 Map Objects 组件式 GIS 开发平台，旋即采用 ActiveX 等控件技术方法建立 Web GIS 软件系统 Internet Map Server for Map Objects，就是一个例子。Map Objects 包含了大量丰富的可编程组件，拥有很强的 GIS 功能和制图功能，用户可通过调用这些构件来建立自己的 Web GIS 应用系统。也就是说，开发者可以使用诸如 VBScript 或 JavaScript 脚本语言、VC、Delphi 和 PowerBuilder 等多种开发工具进行开发。采用这种技术方法构建的 Web GIS 系统，具有很好的灵活性，使 Web 页面成为一种动态的，跳跃的页面，扩展能力强，可充分利用客户机/服务器体系结构优势。这也是现今较为流行的一种 Web GIS 系统实现方法。

6. 利用 Java 编程语言建立 Web GIS 系统

美国 Sun 公司与 1995 年推出的 Java 编程语言是基于网络应用开发的，面向对象的新一代计算机编程语言。它具有对象封装、多态性、继承性等面向对象语言的基本特征，且绝大部分数据类型都是以对象形式出现。由于 Java 无全程和主函数，它的封装性很强。由于采用了虚拟机（Java Virtual Machine）技术，Java 语言程序实现了目标代码与平台无关的特性。同时具有支持 Internet/Intranet 网络模式下的数据分布与计算分布的特性。为保证这些特性的实现，Java 语言通过提供 URL（universal resource locator）对象方式，访问具有 URL 的数据对象，实现数据分布管理；通过将 Java Applet 小程序传送到客户端并下载，实现

计算分布管理。Java语言既可以将计算或其他操作全部在服务器上实现，也可以将部分计算和操作在服务器上完成，其他功能在客户机上运行。由于Java语言所具有的这些功能特点，使其成为实现Web GIS的分布式应用体系结构理想的开发语言。

目前利用Java编程语言开发Web GIS系统可分为两种方式。一种是仅利用Java语言开发客户端的GIS功能，服务器后台仍以传统的开发方式进行或直接对原有的系统进行适当的改造。这种方法的特点是系统开发简单易行，能充分利用原有基础，可以大大缩短系统的开发周期，同时又能保证开发的Web GIS系统具有较强的制图和地理空间分析能力。另一种方式是在客户端和服务器端都采用Java编程语言从系统的底层开发，即简单的GIS功能在客户端解决，较复杂的GIS功能在服务器端完成。这是一种较理想的开发方式，采用这种方法构建的Web GIS具有真正意义上的Browser/Server结构体系。但由于这种开发方式一切都得从底层做起，系统开发的工作量很大，周期长，对一般开发单位具有相当的难度。

上面所述的5种现有Web GIS的工作模式和主要运行环境等之间的比较，如表7-2所示。

表7-2 万维网地理信息系统实现方法比较

类型	工作模式	运行环境	优点	缺陷
基于CGI的Web GIS	CGI	服务器	客户端很小；充分利用服务器的资源	JPEG和GIF是客户端操作的唯一形式，互联网和服务器的负担重
基于服务器API的Web GIS	服务器API	服务器	客户端很小，充分利用服务器资源；速度较快	JPEG和GIF是客户端操作的唯一形式，依附于特定的服务器和计算机平台
基于Plug-in的Web GIS	Plug-in	客户机	具有动态代码的模块；比HTML更灵活，可直接操作GIS数据	与平台和操作系统相关，不同的GIS数据需要不同的Plug-in支持；Plug-in必须安装在客户机的硬盘上

续表

类型	工作模式	运行环境	优点	缺陷
基于 Java Applet 的 Web GIS	Java Applet	服务器 客户机	在支持 Jave 的互联网浏览器上运行，与平台和操作系统无关，分布式处理数据对象	对于处理较大的 GIS 分析任务的能力有限，GIS 数据的保存、分析结果的存贮和网络资源的使用能力有限
基于 ActiveX 的 Web GIS	ActiveX 控件	客户机	具有动态代码的模块，通过 OLE 与其他程序、模块和互联网通讯；是一种通用的部件	ActiveX 需要下载和安装，占用硬盘空间，与平台和操作系统相关；不同的 GIS 数据需要不同的 ActiveX 控件支持

7.1.4 主要 Web GIS 系统平台基本功能

当前，地理信息系统的软件商纷纷推出或升级已有 Web GIS 系统的版本，而且更新的周期越来越快。尤其是几大主流 Web GIS 平台系统提供商，几乎一年左右就有新的版本推出。其中国外比较有代表性的系统平台有 ESRI 公司的 ArcIMS，MapInfo 公司的 MapXtreme，Autodesk 公司的 MapGuide，Intergraph 公司的 GeoMedia Web Map 等。国产 Web GIS 平台有由武汉吉奥信息工程技术有限公司的 GeoSurf，国家遥感应用工程技术研究中心的地网 GeoBeans 等平台。

1. ESRI 公司的 Internet Map Server（ArcIMS）平台

ESRI 于 1995 年开始研发用于 Internet 网上的 GIS 制图软件系统，首先提供给用户的是两个产品平台上的 Internet GIS 制图服务软件：ArcView Internet Map Server 及 MapObjects Internet Map Server 扩展模块。2001 年 ESRI 推出了基于 Internet 和 Intranet 环境的新一代 GIS 应用和地图数据服务产品 ArcIMS。

在 ESRI 的 Internet Map Server（IMS）技术方案中，采用了网络上叫做 CGI（common gateway interface）的技术，对用户的请求在 Web 服务器上运行一个副本用来接收用户请求，并把结果动态分发给用户。具体地在 IMS 里，用户对地理信息的交互空间操作（如用鼠标指定感兴趣的范围或到达的城市）和属性操作（如指定三星级以上的宾馆）可通过 Internet 传送到 Web 服务器上、Web 服务器上的服务程序启动相应的 CGI 程序来解释，完成这种请求，并把结果经由 Web 服务程序发送到 Internet 浏览器上，在浏览器端如同查询本地数据一样。这样就实现了 Internet 上的可视化动态制图技术。ESRI 的 Internet 制图产品支

持灵活的硬件体系结构,Internet服务和IMS服务可在一台计算机上,也可放在不同的计算机上,两者通过TCP/IP通讯。一个Internet服务器可支持同一或不同计算机上的同样或不同的制图应用,IMS能自动分发来自用户的请求,并对性能自动优化,用户可根据各自的应用和网络的负载情况调整其配制。

ESRI推出的新一代平台ArcIMS包括了互联网GIS及制图产品的所有功能。ArcIMS具备了成熟的客户/服务器体系构架,允许对各层进行直接定制。客户端用户通过定制HTML,JavaScript或调用开放的对象模型,可以建立特定的Internet应用程序。而在服务器端,ArcIMS提供了投影和要素专题化等功能。中间件与其他流行的Internet技术兼容,可以使用ASP、ColdFusion、Microsoft Transaction Server、Active Server Pages等技术进行定制,从而扩展WebGIS网站的功能。

ArcIMS其他的一些重要特征还包括:支持要素数据流方式,支持不同来源的数据整合以及使用一些辅助工具。通过要素数据流方式,ArcIMS不仅能够向客户端传输影像数据,而且还可以传输矢量数据。它还允许进行一些超出制图的功能,如数据的空间叠加,地理分析等。ArcIMS通过要素数据流方式,不仅可以通过网络访问远程数据,而且还可以使用自己本地机上的数据。辅助工具包括客户端的要素编辑工具——EditNotes和共享某些特殊地理信息的工具——MapNotes。ArcIMS与其他IMS产品有所不同的一个特征是,ArcIMS是作为NT服务运行的,这一特征使得许多现有用户可以有效地访问网站。

由于ArcIMS具有在服务器和浏览器或其他客户间建立"数据流"的特点,使得用户可以将本地数据与Internet上的数据结合起来,从而为用户和数据发布者提供在互联网上交互制图及访问地理数据的方式。也就是说,它允许用户嵌入自己定制的GIS服务,使用的数据既可以是本地的也可以是分布的数据集,使传统GIS向分布式GIS转变。遍布全球的代理商可以直接为用户发布数据并提供服务。用户可以将广泛分布的数据和本地数据结合在一起,使不同地区的计算机主机协同工作。这种技术使得全球存贮在GIS数据库中的现有空间信息发挥作用。用户可以在任何时间、任何地点共享和使用彼此的数据。任何人通过一个简单的浏览器界面就可以访问经过复杂的专业GIS分析产生的数据。通过不断提高的访问GIS信息的能力,各种组织和社会作为一个整体将更有效地利用现有的空间数据资源。

ESRI的IMS新的体系结构包括许多应用服务,如制图、查询、地理编码等。ArcIMS包括有三个层次。第一层是广域或局域网内的客户端或用户终端,它是直接或间接与服务器相连的计算机或应用程序,通过HTTP或TCP/IP协议与服务器建立通讯。第二层称为中间件,负责管理客户端和服务器端之间接收、发送请求并作出响应的数据,以达到最佳效率与传输速度。第三层是服务器。

其中ArcIMS空间服务器可被看成软件服务器，管理并处理需要在Internet上发布的地理数据。对地理数据的请求被传送到服务器，服务器代理对所需数据执行请求，生成一幅影像或要素地图，进行查询、地理编码或要素提取等等。最后，产生的应答又通过中间件返回到客户端。

ArcIMS的主要特点表现在以下几个方面。

● 简单的基于向导的界面。ArcIMS在创建、设计、管理GIS站点上提出了一套简便易行的解决方案。它的便捷而强大的管理构架可以方便地帮助开发者完成建立Web服务以及优化站点的工作。通过向导引导，系统能够迅速完成创建地图服务、设置网页、发布和管理站点的工作。

● 强大的智能化的客户端。ArcIMS同时推出HTML和Java两种版本的强大的客户端。ArcIMS Java客户端支持真正的客户/服务器方式，它允许客户直接使用本地缓冲区中的数据。客户端的操作不再必须由服务器来完成，许多操作可以直接在客户端快速高效地执行。

● 地图编辑和地图注释功能。ArcIMS客户端支持动态编辑和标注地图信息，允许用户将编辑结果返回给服务器，并帮助记录下对服务器端数据库所做的改变。地图注释功能还允许在地图上标注符号用来提示其他人的注意或以备今后查看，从而提高了网上GIS数据的交互能力。

● 方便的定制功能。ArcIMS的客户端可以通过工业标准语言，如VBScript和JaveScript，方便地定制以适用于不同的功能需求，为最终用户和希望自己定制界面和应用的开发商提供了方法。

● 高质量的制图显示功能。由于ArcIMS支持矢量数据流，可获得更清晰的屏幕显示效果。此外，由于采用了新的影像边界平滑处理技术（anti-aliasing），提高了栅格数据的显示质量。

● 开放的可伸缩的结构。ArcIMS独特的结构提供了一个开放的灵活的平台。不仅支持简单的Intranet及Internet应用，而且支持一个或多个跨服务器的分布式GIS发布应用，便于扩展原有的网上GIS应用，无须重新构建。此外，ArcIMS与其他流行的Internet技术兼容，如ColdFusion, Microsoft Transaction Server, Active Server Pages等。

2. MapInfo公司的MapXtreme平台

MapXtreme是基于Internet/Intranet的地图应用服务器。它采用标准的TCP/IP协议，通过HTTP进行文档和文件传输，在浏览器端为标准的HTML语言，从而保证了与客户端浏览器的无关性。MapXtreme在客户端提供了两种工作模式，一种是标准的HTML网页的模式，只要任何支持HTML的浏览器都可正常工作，例如IE、Netscape、或UNIX平台的浏览器。推荐在Internet上采用这种工作模式。另一种是Java Applet插件，这种方式能够增强在浏览器端的交互

性。但对网络速率要求较高，一般在Intranet上采用。MapXtreme向用户提供Java Applet的源码，便于用户添加和维护自己的应用。在ASP(Active Server Page)环境下，MapXtreme在服务器端的开发语言为VBScript或者JavaScript，在客户端可方便的扩展HTML，Java或者JavaScript支持。选择VBScript的原因，一方面是因为Visual InterDev提供了非常易用的可视化开发环境，而且所采用的编程语言很容易掌握，从而保证了开发周期能够得以控制；另一方面，在CGI程序中对HTTP协议的面向无连接的特性使用非常不便，在编程时往往需要在HTML中用大量的隐藏变量来记录状态参数，以便下一次CGI程序再度被激活时用来作初始化。

MapXtreme支持分布式服务体系结构，能与其他标准的Web服务器相连，因而具有良好的开放性。MapXtreme与Web Server的连接是通过应用服务器完成的，如Microsoft的ASP，以及流行的Haht hahtsite。MapXtreme自带Haht hahtsite开发环境，用户无须另外单独配置网络应用开发工具。一个Web Server可以任意挂接多个MapXtreme地图应用服务器。MapXtreme的服务器可以自动维护和协调Web服务器和多个MapXtreme之间的请求响应关系，无须用户编程解决。MapXtreme采用三层结构，包括客户机、客户机/服务器以及服务器。客户机具有用户接口，进行数据的显示，客户机/服务器负责应用处理过程，服务器端只进行数据的管理工作。这种体系结构使得应用系统能够在客户机和服务器端实现共享，或者运行在一些中间平台，一般称之为"应用服务器"。应用服务器能够进行大量的数据分析工作，因此减少了网络的阻塞。在MapXtreme的工作方式下，所有的地图数据和应用程序都放在服务器端，客户端只是提出请求，所有的响应都在服务器端完成，只需在服务器端进行系统维护即可，客户端无须任何维护，大大降低了系统的工作量。MapXtreme具有的地图化基本功能如表7-3所示。

表7-3 MapXtreme地图化基本功能

专题图	利用晕渲、等级符号、独立值、点密度、饼图、直方图进行区域值的显示
对象处理	合并、缓冲区、相交、删除对象(点，线，面)、返回结果数据
对象编辑	生成、修改、删除
绘制图层	允许开发人员绘制定制的地图对象，例如尺标、天线传送方向的箭头
查找	通过行政名、ZIP码、城市名、街道名或客户进行查找
图层控制	允许用户管理多层地理信息，诸如层的颜色、缩放、可视和层的风格
空间选择	允许用户在规定的矩形、半径和多边形范围内进行选择和操作
数据源	使用通用的数据界面，ODBC、DAO、ClipBoard和OLE Data访问数据

MapXtreme for NT 版以 MapX 为图形内核，将矢量地图转化成 GIF 或 JPG 格式的栅格图像，使用户可以通过 WWW 浏览器访问地图。同时 MapXtreme for NT 版还提供 Java 或 ActiveX 的 Widget，实现多平台上的地图缩放、平移等操作。由于传递到浏览器端的只是一幅经过高度压缩的栅格地图，而真正的矢量地图及数据仍保留在服务器端，因此减少了网络传输负担，同时降低了原始数据被盗用的可能。

MapXtreme for NT 版的技术特点主要表现在以下几个方面

• 地图显示查询。通过 MapXtreme for NT 版，用户可以在 Internet/Intranet WWW 上发布基于电子地图的应用系统。所有的最终用户只需在自己的机器上安装浏览器（如 Microsoft Internet Explorer 或 Netscape）即可访问存放在服务器端的空间数据，用户可以很方便地对地图进行放大、缩小、漫游、查询、统计、图层控制、空间选择等操作。

• 稳定可靠的 GIS 高级管理。MapXtreme for NT 版提供了较全面的地图分析功能，包括专题地图、缓冲区分析、地图目标查找、地图编辑、涂抹层、地理编码、扩展地图库、示例数据等。可直接读取 Lotus Notes、空间选择、访问各种数据源等。例如，访问运行在 Oracle、SQL Server、Informix 上的 MapInfo SpatialWare 的图形数据。

• 中心式软件运行和数据管理。通过服务器进行数据、软件集中管理，使用户的使用成本降低。为了实现高性能、集中化和安全性，MapXtreme for NT 版与 MapInfo 公司的空间数据库引擎软件 SpatialWare 兼容，将地图数据存贮到大型关系库 RDBMS 中。

• 地图应用易于扩充。当应用需求增加时只需通过增加服务器上的用户数即可完成。MapXtreme for NT 版的开放式结构可以在各类 Web 服务器或 Web 浏览器上工作，并且可以利用 ISAPI、NSAPI 以及 CGI Gateway 各自的优势。MapXtreme for NT 版不需要任何 plug-ins，因此可以通过任何 PC 及 Unix 工作站的浏览器访问。

• 以 MapX 为地图引擎。MapXtreme for NT 版以 MapX 为图形引擎。MapX 是一个可编程的 OCX 控件，是可重复利用的可编程对象，它提供绝大部分 MapInfo Professional 支持的地图功能，可以利用编程平台所提供的数据库访问机制，也可以利用自身提供的 ODBC 接口，并可进行数据的智能绑定，在客户端安装并可在授权范围内分发，这是一种较为方便、简洁、经济的桌面地图应用方式，使更多的用户得到廉价的 MapInfo 地图信息技术。

• 方便的开发方式。随 MapXtreme for NT 版提供的开发环境是由 Microsoft 公司提供的 Visual InterDev 开发工具。Visual InterDev 是高度集成化的 Internet 开发环境，开发人员可以利用它可视化地创建并维护 HTML 文档，在

WWW应用系统中集成高级应用逻辑，并管理整个Web应用开发过程。Visual InterDev提供的开发语言VB Script已经早已为广大的开发人员所熟悉，这也为MapXtreme的应用提供了良好的开发基础。从技术角度出发，在Visual InterDev上的开发过程相对于过去比较典型的CGI程序开发过程要简单得多。由Visual InterDev创建的ASP在IIS上运行时，能够自动为每一个客户端维持状态参数。这个特征将使开发人员的工作量大大减轻

• MapXtreme for NT版的优势。使用MapXtreme for NT版，开发人员能集中地控制和维护地图和数据库数据，并集中实现应用程序功能，避免了以往系统的维护、同步困难的问题，尤其适合信息量大，用户多的单位的实际情况。另外，由于使用Web浏览器作为客户端，更使开发人员可以将地图信息系统紧密地与其他系统结合，给用户提供统一完整的综合信息系统。

3. Autodesk公司的Autodesk MapGuide平台

Autodesk MapGuide是Autodesk公司发布的面向Internet和Intranet的Web GIS平台。它以易于操作、应用简便和地图交互能力强等优点而被广泛采用。Autodesk MapGuide具有较强的开发环境，它提供了大量的API函数和灵活的地图制作工具。可以使用Java、Visual Basic、Cold Fusion及其他一些符合Web标准的数据库和报表软件与Autodesk MapGuide集成。同时，Autodesk MapGuide还提供了服务器管理工具，可方便地对应用系统进行维护和管理。Autodesk MapGuide包括下列四个软件产品：

• Autodesk MapGuide Viewer。这是安装在客户端的浏览器，以访问集成的Web站点设计工具。MapGuide Viewer从MapGuide的网站下载并安装到客户端的浏览器上，负责在浏览器中解释传输过来的矢量地图窗口文件(以mwf为后缀)。利用这个插件，用户可以浏览含有mwf格式的地图文件，并对显示的地图进行各种操作，如放大缩小、漫游、测量距离、查看标注及制作属性报表、根据地名查询等；也可采取多种选择方法在地图上选取目标；利用分层、分类开关来显示图层；自动按照比例尺显示、消除目标；将浏览到的地图粘贴或打印；按照指明的URL跳到其他页面上去浏览相关资料。

• Autodesk MapGuide Author。MapGuide Author是在Web站点上创建、修改和发布基于矢量交互地图的工具，地图被分成点、线、面、文字4种要素类型的图层，每层数据是一种要素类型的集合。通过层的分类组织可以加快显示速度，降低数据的网络流量。MapGuide Author提供"傻瓜式"面向对象的编程，用户通过设置对话框中的操作属性就可以完成所有制作。MapGuide Author可使地图与数据库相关联，内置SQL查询。通过MapGuide Author可以设置地图访问权限，增加安全性措施，如可通过设置密码、到期自动失效、限制访问次数等方法对用户的访问范围及访问内容加以规定。

• Autodesk MapGuide Server。MapGuide Server 是运行于 Windows NT 上的与 Web 服务器进行交互信息的地图 CGI 服务器软件。它是 32 位多线程软件，可并发连接、访问位于站点上的多个地图文件和关系数据库。MapGuide 有自己的矢量地图格式，以 sdf 为后缀，其他文件格式如 mif、shp、csv、bna 等均需转换成这种格式。由于结合了 Windows NT，可以实现查看事件日志、错误日志、访问日志等功能。

• Cold Fusion。Cold Fusion 用于提供快速 Web 应用开发，包括用于向浏览者发布空间数据。

Autodesk MapGuide 既可采用"无插件"方式通过服务器端 Java 小程序向任何客户端浏览器发送地图和各种图形，实现对图形的平移、缩放、查询；又可通过小巧的客户端浏览器插件(plug-in)提供更多更强的地图交互功能。Autodesk MapGuide 支持几乎所有常见的栅格图形数据格式，如 GIF、TGA、CALS、PNG、FLIC、PCX、BMP、JPEG、TIFF/GeoTIFF、GeoSPOT/BIL 等。

Autodesk MapGuide 起用一个主控文件 Map Window File(MWF)和一个数据文件格式(SDF)，以优化在万维网分布式体系结构上的性能。MWF 文件是由 Autodesk MapGuide Author 软件创建的，一方面存贮指针，指向二维矢量数据文件(Autodesk MapGuide SDF 文件)，另一方面存贮属性关联数据(包括目标数据和在外部数据库中存放的数据的联接，比如，数据库 Oracle，MS Access 等)，目的是提高 Autodesk MapGuide Viewer 在 Intranet 或 Internet 上的访问查询速度。Autodesk MapGuide Release 支持的矢量文件格式包括：ESRI (ArcView) SHP、ESRI Arc/Info coverage、Intergraph DGN、MapInfo MIF/MID、Atlas、BNA、CSV (comma delimited file)、Autodesk MapGuide SDF(Spatial Data File)和 SDL 文件、DWG 文件等。通过将这些文件转换成 SDF，提供给 MWF 使用。

新的 SDP(Spatial Data Provider)结构支持数据扩展，使 Autodesk MapGuide 能够直接使用不同 GIS、CAD 和关系型空间数据库原有的空间和属性数据，如 DWG(AutoCAD Map 2000 及以前版本)、SHP、Oracle8i Spatial、Autodesk VISION 等。也就是说，地图和图形数据，不需进行特殊的转换。此外，新的符号、线型及文本支持，保证了绘图质量，使在线显示如同在 GIS 和 CAD 系统中一样清晰。

4. Intergraph 公司的 GeoMedia Web Map 平台

国际著名 GIS 软件商 Intergraph 也是最早推出 Web GIS 产品的公司之一。Intergraph 公司的 Web GIS 平台 GeoMedia Web Map 采用了分布式体系结构，以 HTML 为应用层协议标准，以通用的浏览器为客户端软件，采用微软 Active Server Pages(ASP)技术，通过客户端浏览器向服务器提出申请，所有的程序都在服务器端执行。当程序执行完毕后，服务器仅将执行的结果返回给客户浏览器。GeoMedia Web Map 运用了超图空间数据仓库技术，可以直接访问商业关系

数据库中的地理空间信息，发布动态的GIS页面，提供矢量地图的发布、检索、放大缩小、漫游、动态视窗等功能。GeoMedia Web Map的技术特点主要表现在以下诸多方面。

• 直接访问多源数据。GeoMedia Web Map主要的优点是，它利用了Intergraph独特的地理数据服务器和超图空间数据仓库技术，能够对多源数据进行无缝集成，可以直接访问并发布多种格式的GIS数据，无须转换。当与数据服务器技术相配合时，GeoMedia Web Map可将来自多种空间数据源的各类不同地理数据库结合起来，这些多种空间数据源包括MGE、FRAMME和CAD文件，以及ArcView shapefiles、Arc/Info coverages、MapInfo、Genamap、Small World和Oracle SC/SDO、SQL Server、Access、ODBC Tabular数据等。能同时处理多种不同格式的空间数据的能力使得Intergraph能为Open GIS新概念的倡导者和先行者。

• 动态发布Active CGM矢量图形。GeoMedia Web Map发布的图形方式是国际标准的开放式矢量数据格式ActiveCGM，动态地发布矢量图形，传送数据量小，传送速度快，便于用户获取数据，并进行查询、分析。

• 实时发布实时更新。GeoMedia Web Map可以动态地发布直接来自最新操作中的GIS数据库中的数据，这些数据包括以矢量为基础的图形、无须将其转成新的格式，也可对此数据库进行复制。GeoMedia Web Map所发布的信息可以做到实时、动态，与数据库的更新保持同步。无须或需要很少的专业软件就能生成供发布所需的图形数据，用户可以通过显示和查询功能就能生成自己的交互式图形，通过Intranet或Internet浏览和检索到空间数据信息。

• 实现客户端空间分析。GeoMedia Web Map提供了各种基于Server的分析组件，用户在客户端只需通过浏览器就可以进行各种专业GIS分析，如路径分析、缓冲区分析、地理编码、标注、空间分析等。

• 易于开发。使用GeoMedia Web Map发布空间图形信息需要很少或无需编程技能。使用标准的Web开发工具，例如Java Script、VB Script、FrontPage和ActiveX等，GeoMedia Web Map就可被用户化。GeoMedia Web Map的技术核心是GeoMedia。由于GeoMedia的设计是基于Windows 95/98/2000和Windows NT风格的，它的操作一般用户都很熟悉，就像其他基于Windows的工具一样。这样使得在Windows的环境下，非技术人员操作GeoMedia非常容易。同时GeoMedia Web Map提供了大量的控件和对象，用于二次开发，以完成基于Web的各种GIS空间分析应用。

• 维护简单。由于GeoMedia Web Map直接发布数据库中随时更新的数据，发布时也不必转换为其他格式，所以对于所发布信息的更新和维护非常简单。另外，ASP的源程序在服务器端，不会被传到客户浏览器，因而可以避免所写的源程序被他人剽窃，也提高了程序的安全性。

• 客户端操作简便。在客户端，使用者不必具备任何 GIS 知识，只需浏览器即可完成对 GeoMedia Web Map 发布信息的浏览、查询和分析。GeoMedia Web Map 支持基于不同操作系统环境（如 Windows 9x、Macintosh、UNIX）的多种浏览器，如 Microsoft Internet Explorer 和 Netscape Navigator 等。

• 栅格影像背景。GeoMedia Web Map 允许使用栅格影像作背景（二值或是连续色调），支持大数据量的镶嵌影像。使用 Intergraph 公司拥有的先进图像浏览技术，影像数据量可以保持到较小程度，优化了 Web 上的反应时间。

• 多媒体支持。GeoMedia Web Map 遵循 Web 开发工业标准，能将光栅图像、视频、音频与 GIS 较好地集成。这些性能扩展了可视化 GIS 的能力。通过标准的 GIS 技术，GeoMedia Web Map 可将其特征与其他多媒体信息相连，生成超级连接显示和表达。这些能力使得 GeoMedia Web Map 成为极具竞争力的 Web GIS 开发平台。

5. 武汉吉奥公司的 GeoSurf 平台

GeoSurf 是由武汉吉奥信息工程技术有限公司经过三年努力于 1999 年底推出的国产 Web GIS 软件。GeoSurf 基于 Internet/Intranet 的分布式计算环境，系统模块采用了组件化构造和分布式处理方式，同时以 Java/JDBC 构造多数据源地理信息互操作中间件。GeoSurf 将矢量图形与数据库无缝连接，与硬件环境无关，实现了异质数据的透明获取、操作。GeoSurf 还包括了一个后台 GeoAdmin 运行模块。Geosurf 用于客户端矢量图形的放大、缩小、漫游、查询以及专题图制作、图形输出等一系列 GIS 信息的远程访问，GeoAdmin 是针对 GIS 地图服务可视化而开发的服务器端软件环境。GeoAdmin 通过为 GIS 数据的在线发布提供可视、自动、持久、安全的管理机制，实现空间数据多源集成、图幅自动定制、系统安全性管理、地名索引表创建、多媒体数据结合以及 GIS 的数据编辑等等功能。

Geosurf 主要技术特点和功能包括：适合于客户端不同操作系统平台，通过目录服务实现分布式数据源的组织、管理以及图形分层调用和显示，使用地名数据库直接查询地名信息并浏览对应的矢量图形，实时生成基于矢量的点线符号和智能化显示；常用地图及专题图制作、多层叠加、图符拖动、风格修改、图例、统计图（三维旋转直方图、曲线图、饼图）等图形功能；可视化方式下图形与属性互查、最佳路径选取（以路况、路长、交通状态为判断因子）、可视化操作多种比例尺图形调用和生成、地理数据属性报表生成打印；基于超图模型的声音、文本、图像、视频、虚拟现实与空间信息的交互、直接读取多种数据源数据（Arc/Info、MapInfo、MGE、GeoStar、AutoCAD 等）；大型关系数据库管理海量属性数据多种语言（汉、英、日等）实时转换等。

GeoAdmin 的主要功能包括：数据加密传输、图层加密管理、用户访问权限

控制；可视化方式组织、管理、发布空间数据；二次开发定制个性化界面；实现地图定制、属性编辑、点线符号自动添加、不同比例尺图幅连接及逐级放大。

GeoSurf 支持的矢量数据格式有：内部格式 GeoFile 和 GeoDB、MapInfo 的 MIF/MID、MGE 的外部交换格式、ARC/Info 的 Coverage/Shape 和 E00 格式、AutoCAD 的 DXF 格式等。

6. 国家遥感应用工程技术研究中心的地网 GeoBeans 平台

地网 GeoBeans 是由国家遥感应用工程技术研究中心开发的 Internet/Intranet 分布式计算环境网络 GIS 平台，采用 Browser/Server 结构，提供了网络化解决方案。地网 GeoBeans 使用简单、维护方便，具有支持二次开发、跨平台运行等特性。地网 GeoBeans 采用与平台无关的 Java 语言 JavaBeans 构件模型，可在多种系统平台上运行（WINDOWS 或 UNIX），支持栅格和矢量数据处理。地网 GeoBeans 的主要功能包括：空间矢量图形的分层显示和管理、栅格图形与矢量图形的放大、缩小、平移等操作；兼容多种空间数据格式，可以读取多种数据格式（GBD、ArcInfo、ArcView、DXF、Map GIS、MapInfo 等）；采用了 JDBC 技术，可以与多种大型数据库（Oracle、Sybase、SQL Server）连接；空间、属性数据双向可视化查询、统计数据图形化分析（三维直方图、曲线图、饼图等）；路径分析、（乘车路线、最佳路径、管网管理、公交站点查询等）、用户动态制图及个性化地图输出；兼容 MIS 系统、智能二次开发向导、并发处理多用户。

地网 GeoBeans 的特点是使用及维护简单，用户无需具备复杂的专业知识，便可在很短的时间内安装和配置软件，构建 GIS 系统。针对网络的不同状况，地网 GeoBeans 可将应用分布在不同的机器上。由于采用了 Java 的 RMI 技术，实现了对象级的分布式计算的互操作。

地网 GeoBeans 包括服务器端图形转换、图形编辑、符号编辑、图形管理、空间分析、三维可视化、智能发布向导和发布服务器八个模块及客户端 CMExpress 模块。

- 图形转换模块。能够读入的数据格式有：ArcView、Shape、Arc/Info E00 交换格式，MAPGIS 明码格式，AutoCAD DXF 文本格式，MapInfo mif 格式，Raster Data TM 格式，GIF 图像格式等。
- 图形编辑模块。主要功能有：点、线、面、网络等空间图形编辑，属性数据编辑，拓扑关系生成，空间分析（特定专题图的制作）等。
- 符号编辑模块。主要是编辑 GeoBeans 使用的点矢量符号和线型符号，并可对符号进行分类管理，关键词搜索。
- 图形管理模块。主要功能有：浏览及检查数据的正确性，采用图幅方式按工程组织发布数据（空间与属性），以图库格式压缩输出数据，空间数据和属性数据一体化管理，具有部分空间计算功能。

• 空间分析模块。提供对空间数据的缓冲区分析、叠加分析、等值线分析（空间插值）等空间分析功能。

• 三维可视化模块。主要提供三维表面生成，三维景观生成，三维场景制作，三维表面分析和空间查询功能，以满足数字地面模型、数字城市制作及其应用工程的需要。

• 智能发布向导模块 GeoBeans智能发布向导分为智能栅格发布向导和智能矢量发布向导。主要有：显示并定位GeoBeans系统的安装目录；指定要发布的地址和工程文件；设定发布名称和工程描述等功能。

• 空间信息发布服务器（MapServer）。本模块的主要功能是管理地图服务的配置和装载、在线监控客户端的服务请求、维护服务器的系统日志文件等。

• 客户端模块CMExpress。GeoBeans客户端图形浏览器模块，统称为地图浏览器CMExpress，或称为客户端模块CMExpress。根据功能的不同划分为三种型号，依次为A型，B型和C型，分别位于三个文件中，ncgmea.jar（A型）、ncgmeb.jar（B型）、ncgmec.jar（C型）。地图浏览器CMExpress以Java软件包形式提供用户使用。用户能接触到的是它的外部特性：属性、事件和方法。

§7.2 虚拟现实地理信息系统

虚拟现实地理信息系统（virtual reality GIS，VR GIS）是虚拟现实（virtual reality，简称VR）与地理信息系统相结合的技术，是一种用于研究地球科学的，或以地球科学系统为对象的虚拟现实技术，也是近年来一个新兴的地理信息系统研究分支。虚拟现实和地理信息系统的结合，一方面有利于充分体现GIS空间表达的优势，提高空间表达、分析和查询的能力，另一方面扩大了虚拟现实的应用领域，使之更好的与各种地学和区域科学结合并真正为之服务。近年来，VR GIS已开始应用于城市规划、水利防汛、流域治理、地质构造分析和环境治理规划等多种领域。

7.2.1 虚拟现实技术

1. 虚拟现实技术基本概念

虚拟现实技术，有时称为实时环境、虚拟空间、人造现实、仿真技术等等。在计算机图形研究中，也常常称作"灵境"。它是近年来十分活跃的一个领域，是包括计算机图形学、多媒体技术、人工智能、人工接口技术、高度并行实时计算技术和人类行为学研究等一系列高新技术的汇集，以及这些技术在更高层次上的集成和渗透。它给用户以更逼真的体验，为人们探索宏观世界和微观世界提供了极大的便利，是多媒体技术发展的更高境界。虚拟现实自诞生以来，就

在航空航天、军事、核工业以及其他行业中发挥着不可替代的作用，与多媒体、网络并称为三大前景最好的计算机技术。

虚拟现实(virtual reality)的定义目前仍不统一。一般说来，虚拟现实是指运用计算机技术生成一个逼真的，具有视觉、听觉、触觉等效果的可交互、动态世界，人们可以对该虚拟世界中的虚拟实体进行操纵和考察；是一种让用户在人工合成的环境中获得"进入角色"的体验。

2. 虚拟现实技术的主要特征

虚拟现实有三个突出特征，也是人们熟称的"3I"特征——交互性(interactivity)、沉浸感(illusion of immersion)和想象(imagination)。这三个特征可以用来区别相邻技术，如多媒体技术、科学计算可视化技术等。交互性主要指参与者通过使用专用设备，用人类的自然技能实现对模拟环境的考察与操作的程度，例如用户可以用手去直接抓取模拟环境中的物体，而且用户有抓取东西的感觉，还可感觉到物体的重量，视场中被抓取的东西也应立刻随着手的移动而移动。由于VR并不只是一种媒体或一个高级终端用户界面，它的应用能解决在工程、医学、军事等方面的一些问题，这些应用是VR与设计者并行操作，这依赖于人的想象力，由此而产生VR的第三个特征——想象。而VR最主要的技术特征是沉浸感，也有人称为"临境感"、"投入"、或"存在感"。VR力图使用户在计算机所创建的三维虚拟环境中处于一种"全身投入"的感觉状态，有身临其境的感觉，觉得自己是该虚拟环境中的一部分，而不是旁观者。

3. 虚拟环境概念的四层含义

虚拟现实是由计算机和电子技术创造的世界，是一个看似真实的模拟环境，通过多种传感设备，用户可根据自身的感觉，使用人的自然技能对虚拟世界中的物体进行考察和操作，参与其中的事件；同时提供视、听、摸等直观而又自然的实时感知，并使参与者"沉浸"于模拟环境中。虚拟环境的概念包括了以下四层含义。

第一，"模拟环境"，就是由计算机生成的具有双视点的实时动态的三维立体逼真图像。逼真就是要达到三维视觉，甚至包括三维听觉、触觉、嗅觉等的逼真；"模拟环境"可以是某一特定现实世界的真实再现，也可以是虚拟构想的世界。

第二，"感知"，是指理想的虚拟现实应具有人类所具有的感知，除了计算机图形技术所生成的视觉感知外，还有听觉、触觉、力觉、运动等感知，甚至还包括嗅觉和味觉等，也称为多感知(multi-sensation)。由于相关技术受到传感器的限制，目前所具有的感知功能仅限于视觉、听觉、触觉、力觉、运动等，嗅觉方面也有了新的进展。但无论是感知的范围还是精确程度都无法与人相比拟。

第三，"自然技能"，指的是人的头部运动、眼睛、手势或其他人体的行为动

作。由计算机来处理与参与者的动作相适应的数据，对用户的输入（手势、口头命令等）作出实时响应，并分别反馈到用户的五官，使用户有身临其境的感觉。用户成为该模拟环境中的一个内部参与者，还可以与该环境中的其他参与者相互交流。

第四，"传感设备"，指三维交互设备，常用的有立体头盔、数据手套、三维鼠标、数据衣等穿戴于用户身上的装置和设置于现实环境中的传感装置（不直接戴在身上），如摄像机、地板压力传感器等。

需要指出的是，VR 并不是真实的世界，也不是现实，而是一种可交替更迭的环境，人们可以通过计算机的各种媒体进入该环境，并与之交互；它是在众多的相关技术（如计算机图形学、仿真技术、多媒体技术、传感器技术、人工智能等）基础上发展起来的，但又不是这些相关技术的简单结合。从技术看，VR 与各相关技术有着或多或少的相似之处，但在思维方式上看，VR 已经有了质的飞跃。它是一门系统性的技术，不能只从某一个方面考虑问题，而需要将所有组成部分作为一个整体去追求系统整体性能的最优化。它把抽象、复杂的计算机数据空间表示为直观的用户能感知到的事物，它的技术实质在于提供了一种高级的人与计算机交互的接口。

7.2.2 VR GIS 基本概念及其特点

虚拟现实通常是抽象环境的可视化，而 VR GIS 则是在"真实的"背景上对"抽象"对象的混合表示。目前情况下，VR 技术和 GIS 技术两者的连接，主要是通过虚拟现实建模语言（virtual reality modeling language，简称 VRML），把 GIS 信息转到 VR 模块中表示。VR GIS 方法是基于一个耦合的系统，由一个 GIS 模块和一个 VR 模块组成。现在较多的 VR GIS 是构建在专业 GIS 及遥感平台之上或运用虚拟现实建模语言来实现的。VR GIS 所采用的数据库是传统 GIS 数据库，既可以虚拟过去和未来的地理环境，使用户可以在空间维上进行漫游、查询、分析，还可以表示时间维上的信息。VR GIS 使传统的二维向三维方向发展迈出了重要的一步。

1. VR 技术与 GIS 技术的结合

当前，GIS 技术和虚拟现实技术各自都在迅速发展之中。自 20 世纪 90 年代以来，GIS 研究开发人员开始将虚拟现实技术引入 GIS 中，目前虽有不少单位、组织和个人从事这方面的研究，但所开发的 GIS 和虚拟现实相结合的应用系统，大多还处于虚拟现实与 GIS 数据结合的层次上，即将虚拟现实系统作为用户与 GIS 系统交互的界面，而将 GIS 作为空间数据库管理系统。但在两者数据结合方式上，虚拟现实技术和 GIS 技术已开始从松散的数据转换发展到紧密的系统集成上来。

虚拟现实和GIS的松散数据结合，是指利用公用数据接口，通过动态数据转换器在虚拟现实系统和GIS之间进行有限的信息交换，用户在虚拟现实系统中执行的操作被翻译成对GIS空间数据库的操作，然后传送到GIS中，由GIS完成，再将操作结果转换到虚拟现实系统中提交给用户，如图7 2所示

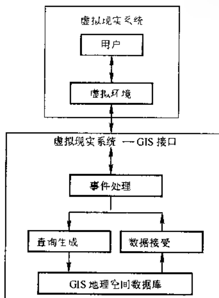


图7-2 VR和GIS系统的松散数据结合

目前，VR GIS系统的数据库仍采用传统的GIS数据库，VR的功能主要用于增加GIS的制图显示功能，并且越来越多的解决方案采用VRML标准。随着互联网技术的发展，VR GIS已开始向分布式方向发展，具有了Internet的一些基本功能，如在浏览器中动态显示三维地表景观等。

实现虚拟现实与GIS在数据模型和功能上的统一，是当前VR-GIS发展的重要方向，虚拟现实建模语言（VRML）标准和OpenGIS（Open geodata interoperability specification，开放性地学数据互操作使用规范）数据模型和服务模型的建立，使两者在数据模型和功能上的统一成为可能。如图7 3所示，VR和GIS集成在一个整体系统中。

将模型从GIS平台环境移植到虚拟环境中，将会有利于用户更方便的分析空间数据。目前这方面的研究工作，主要是将GIS中与栅格有关的应用模型集成到虚拟环境中，如火灾模型、污染扩散模型等。

VR GIS技术目前还不用数字化头盔、手套等设备，运用VRML技术，可以

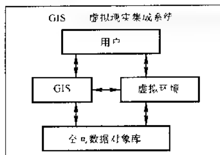


图7-3 VR和GIS的集成

在微机上进行，使费用大幅度降低，所以具有被广大用户接受的特点。但是，当前大多数VR GIS只具有三维立体、动态、声响，即具有视觉、听觉、运动感觉的特点，没有触觉，更没有嗅觉和可操作特点等。不过通过大脑的联想，也可以产生一定程度的身临其境的感觉，如洛杉矶城市改造的虚拟、恐龙及其生长环境的虚拟等。值得注意的是，"侏罗纪公园"、"龙卷风"等娱乐性虚拟影视等典型例子，这些不是真正的虚拟，而是介于虚拟现实与计算机仿真之间的一种过渡技术，是一种准虚拟或不完善的虚拟，或称之为半虚拟技术。

2. VR GIS基本概念及特点

VR GIS在20世纪90年代开始出现，是一种专门用于研究地球科学，或以地球系统为对象的虚拟现实技术或计算机仿真技术或多媒体技术，是虚拟现实与地理信息系统相结合的产物。近年来，VR GIS甚至融入到了万维网地理信息系统(Web GIS)和组件式GIS(ComGIS)之中。VR GIS既具有传统GIS的特点，如数据的存贮、处理、查询、分析功能，又具有VR界面及其交互方式。值得指出的是，VR必须具有三维特征，而当前的GIS软件和数据库多是二维的。因此，GIS向VR GIS的转换并不是透明的。一般认为，一个理想的VR GIS应包括下列特征。

- 对现实的地理区域的非常真实的表达；
- 用户在所选择的地理带(地理范围)内外自由移动；
- 三维(立体)数据库的标准GIS功能(查询、选择、空间分析等)；
- 可视化功能必须是用户接口的自然整体部分。

在VR GIS发展初期，GIS和VR技术的连接，主要是通过VRML进行文件格式转换，把GIS信息转到VR中表示。因此，VR GIS是一个合成的综合系统。当前VR GIS技术发展的特点主要表现在以下几个方面。

- 区域表达的真实性。地理区域表达越真实，越容易沉浸在虚拟的地理环

境中，这正是VR工作者所追求的目标之一，让用户有身临其境的感觉，即"沉浸感"。

• 空间、时间维的漫游、查询 VR GIS可以虚拟过去和未来的地理环境，使用户不但可以在空间维上查询、分析，而且还可以查询时间维上的信息，从而实现时空维上的自由漫游

• 用户和系统之间的交互作用。交互性是VR系统最基本的要求，也是VR GIS最本质的特征之一。用户是交互作用的主体，可以通过手势、口令等多种方式与虚拟的地理环境打交道。

• 信息丰富 VR GIS最重要最基本的数据库是GIS数据库，包含海量的地理信息。此外还有影像、声音等多媒体信息。尤其值得一提的是VR GIS将遥感信息再现到虚拟的地理环境中来，实现遥感和GIS分析的自然统一。

GIS是管理地理空间数据和属性数据的平台。用户和GIS系统的交互主要是通过专业化的界面或通过命令进行，用户难以有效的利用地理信息。而虚拟现实地理信息系统的优势就是营造用户容易接受和理解的多维化信息空间，使用户沉浸于虚拟环境之中，更易于分析、处理信息。虚拟现实和GIS技术的结合，对促进用户与地理信息系统之间的交互，使用户能更加充分的利用和分析地理信息，开发更高层的GIS分析功能具有非常重要的意义。

GIS与虚拟现实的结合是一个循序渐进的过程，大致要经历三个发展阶段

第一阶段，数据的结合。这个阶段是GIS和虚拟现实结合的初级阶段，主要是充分利用GIS技术和虚拟现实已取得的成果，将GIS作为后台空间数据管理的工具，而将虚拟现实作为前台用户与地理空间信息交流的渠道，为用户提供便捷高效的查询、分析功能和结果反馈的途径。

第二阶段，数据模型和功能的统一。在这个阶段，整个系统是建立在统一的系统结构之上的，虚拟现实和GIS采用统一的数据模型、建立统一的空间数据库。用户不但可以利用虚拟现实系统提供的交互能力来观察空间数据，而且可以利用虚拟现实系统交互接口来实现GIS查询、分析等空间数据管理功能。

第三阶段，虚拟现实和GIS技术一体化。这是虚拟现实和GIS结合的高级阶段。考虑到地理空间数据的特点和地学应用领域本身的特点，将不同的地理模型集成到虚拟环境中，让用户在虚拟环境中更直观地操作和调控地学模型，利用虚拟环境提供的各种模型解决许多地学应用问题。这种虚拟地理环境不但可以分析当前地学实际问题，也可以通过虚拟过去和未来某一地理现象的场景，为地学中的若干问题提供科学的论证环境，以更好地为决策者服务。

3. VR GIS与三维地形可视化

VR GIS技术的主要应用之一是对地表形态进行三维动态表达，它使GIS的三维地形可视化进入较高的技术层次。三维地形可视化的基本概念已见于本

书§4.4，这里再作些阐述。

可视化是指在人脑中形成对某物或某人的图像，是一个心智的处理过程（mental process），能促进对事物观察力的提高及概念的建立；可视化实质上是用图形图像来解释数据的一种方式。可视化有时又称为科学可视化（scientific visualization），也有人称为科学计算可视化。科学可视化以计算机图形学和计算几何等为其主要的知识基础。

表面可视化技术是一种重要的计算机可视化技术，地形三维可视化可以看作是表面可视化技术在地形表面上的应用。地形可视化是一个基于 DEM 的显示、简化和仿真内容的学科分支。如前所述，GIS 目前的地形三维可视化本质上是二维的，它将 Z 值投影到一个模型上进行三维显示，实际上并不具有地形下面和内部的信息，因而也称为 2.5 维的"面三维可视化"。目前这种 2.5 维表达的面三维可视化技术已基本成熟，地形可视化已广泛应用于战场环境仿真、娱乐与游戏、地形的穿越飞行、城市规划、土地管理与利用、水利防汛、流域治理、气象预报等领域。

三维建模的主要途径是高程矩阵（Grid）和 TIN。图 7－4 给出 Grid（左）和 TIN（右）两种 DEM 模型的透视立体图。两种方法各有优缺点（§4.4.2）。当要保持原始数据的精确性，或研究区域较小时，一般采用 TIN 结构。当采样点或数据密度足够大时，二者之间的性能差别就变得不太明显。实验结果表明，用 Grid 建立的三维地面模型运行速度明显大于用 TIN 建立的地面模型。因此，在研究较大区域应用中，不要求十分精确或不必准确表现地物要素时，选用 Grid 结构较为适宜。

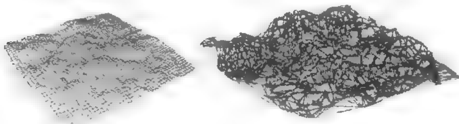


图 7－4　Grid（左）和 TIN（右）模型的透视立体图

VR GIS 技术使 GIS 的三维地形可视化进入较高的技术层次。实现 VR GIS 层次的地形可视化，首先需要逼真显示。目前产生逼真地形显示的方法主要有下述几种：

一是将航空像片或卫星影像数据映射到 DEM 上。这种方法可以逼真地显

示地面各种地物和人工建筑物的颜色纹理特征，而表现地形起伏产生的几何纹理特征不大明显，因而常用来表达地面起伏较平缓、地物丰富和人类活动较频繁地区（如城镇、交通沿线等）的地形。

另一种是用一定的光照模型模拟光线射到地面时所产生的视觉效果，经明暗处理产生具有深度质感的灰度浓淡图像，并用各种方法模拟地形表面的各种微起伏（几何纹理）和颜色纹理，如添加纹理图像和分形纹理（利用分形产生植被、水系等）等。使用这一方法存在着内存与速度的问题、着色算法由于叠加了纹理而复杂化，影响显示速度；而提高显示速度的措施常需要较大的内存支持。例如，可以在多分辨率地形模型中加入多分辨率的纹理，即将图像分成多级分辨率，然后根据视点的变化来选择其中某一级分辨率的图像；或在系统中通过建立金字塔式的图像分辨率，以及采用贴图（drape）方式来叠加纹理等，这都需要较大的内存。

叠加三维地表物体也是逼真地形显示的途径之一。该方法首先必须对叠加的地物，如道路、河流、建筑物、桥梁等进行三维空间造型。目前有许多软件可以用来进行三维建模。如3D MAX、AutoCAD等。对复杂的三维造型，许多商用地形可视化系统（如MultiGen）开发了专门的三维造型生成器，可以先由该三维造型器生成各种地物，然后再贴到地形的表面。此外，还可用一些非商用系统的三维造型工具来建模，然后再导入到三维地形可视化体统中。对于简单的建筑物，可以将二维线形或二维矩形拉伸到一定高度形成三维实体。

逼真地形显示再加上有身临其境的动态漫游，就基本上达到目前的VR GIS层次的地形可视化的要求，或者说达到前面所述的"准虚拟"GIS地形可视化的要求。动态漫游的功能可以基于不少的GIS或图形图像软件平台而获得。

目前已有不少GIS软件开发出相对成熟的VR GIS功能，如美国ESRI公司的ArcView GIS 3D Analyst、Intergraph公司的MTA和ERDAS公司的Imagine Virtual等。基于这些平台，用户可以比较方便地实现地形可视化或VR GIS层次的地形可视化。考虑到这些软件通常比较昂贵和发展国产GIS软件的重要性，我们也可以基于更通用的平台自主开发VR GIS软件。下面就来介绍一种基于OpenGL的开发途径。

7.2.3 三维显示工具OpenGL技术

1. OpenGL基本概念

OpenGL是一种可产生高质量图像的三维图形设计接口，它是由运行在SGI公司的IRIS GL图形工作站上的IRIS GL三维图形接口发展起来的。OpenGL是一个开放的图形标准，它不仅是一个图形库，更是一个API（application programming interface），其本身是一个与硬件无关的编程接口，可以在不同的硬件

平台上得到实现,可运行在SGI公司、DEC公司、IBM公司等几乎所有的图形工作站上。OpenGL(open graphics library)以其多平台性、网络透明性、系统稳定性等卓越性能已成为三维图形制作方法中事实上的工业标准,由独立的ARB(architectural review board)委员会制定和维护。

OpenGL的设计思想是程序设计人员只需定义要画出的场景物体、光照条件、视图参数等,而由OpenGL来完成渲染场景所需的其他一切工作,包括各种坐标变换、投影变换、光照设计、裁剪、隐藏面消隐等。与一般的图形开发工具相比,OpenGL具有以下突出特点。

• 多平台性。OpenGL几乎能在所有主要平台上运行,包括Mac OS、OS/2、UNIX、Windows 95/98、Windows NT/2000、Linux、OPENStep、Python以及BeOS等,也能很好的与绝大多数的窗口系统一起工作,包括Presentation Manager、Win32、X/Windows系统等。

• 网络透明性。建立在客户/服务器模型上的网络透明性是OpenGL固有特性,它允许一个工作站上的进程在本机或通过网络在远程工作站上显示图形。利用这种透明性能够均衡的共同承担图形应用任务在各工作站的负荷,也能使得没有图形能力的服务器能使用图形工具。

• 出色的编程特性。OpenGL在各种平台上已有近十年的使用经验,加上严格的规范控制,使得OpenGL具有良好的稳定性;而且由于OpenGL规范由OpenGL体系结构评审委员会独立负责管理,因此OpenGL具有充分的独立性。

• 高质量和高性能。OpenGL的高质量和高性能使得三维动画、CAD、可视化、仿真等技术的优越性得到最高程度的发挥,使广告、CAD/CAM/CAE、娱乐、医学图像处理、虚拟现实等领域里的开发人员能够创造出质量极高的二维和三维图像。

• 应用广泛。OpenGL是目前最主要的二维和三维交互式图形应用程序开发环境之一,已成为开发应用广泛的图形应用编程接口。自1992年以来,OpenGL已广泛应用于CAD/CAM、三维动画、数字影像处理、三维建模以及地理场景虚拟现实等领域,Kinetix公司的3D Studio VIZ和3D Studio Max就是突出的代表。开发人员可在微机、工作站及超级计算机上使用OpenGL提供的各种功能。

2. OpenGL工作结构

OpenGL包含了一个基本API图形库和其他辅助库,如图7-5所示。这类API的主要功能包括物体描述、旋转、平移、缩放、材质、光照、纹理、像素、位图、文本处理等。"GLU"表示实用API,其主要功能包括绘制二次曲线、NURBS曲线曲面、复杂多边形以及纹理、矩阵管理等。"WGL"是Windows为支持OpenGL而特别设计的编程接口;"GLX"是UNIX系统OpenGL的编程接口。

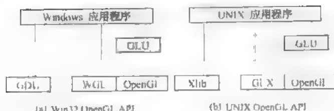


图7-5 OpenGL API结构图

为实现与硬件平台无关，OpenGL不提供窗口管理、输入管理和事件响应机制，因此OpenGL程序必须使用所在平台的用户接口（如GDI和Xlib）。

如图7-6所示，OpenGL指令的解释模型是客户/服务器模式，即客户端采用OpenGL进行绘制工作的应用程序向服务器（OpenGL内核）发布命令，这些命令由服务器来解释。在大多数情况下，客户和服务器是运行在同一台计算机上的。基于客户/服务器模式，在网络环境下不同计算机上多个客户可以得到在其他计算机上服务器的服务，即OpenGL具有网络透明性。

OpenGL的库函数被封装在Opengl32.dll动态链接库中，从客户应用程序发布的对OpenGL函数的调用首先被Opengl32.dll处理，在传给服务器后，被Winsrv.dll进一步处理，然后传递给DDI（device driver interface），最后传递给视频显示驱动程序。

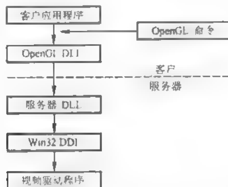


图7-6 客户/服务器模式下的OpenGL运行机制

作为图形硬件的软件接口，OpenGL最主要的工作就是将二维和三维物体绘至帧缓存。这些物体由一系列描述物体几何性质的顶点（vertex）或描述图像的像素（pixel）组成。OpenGL执行一系列操作把这些数据最终转化成像素格式

并在帧缓存中形成最后结果。

图7-7是OpenGL的绘制原理图，用户指令从左侧进入OpenGL。指令分为两部分：一部分画指定的几何物体，另一部分指示在不同的阶段怎样处理几何物体。为加快显示速度，许多指令可能被排列在显示列表(display list)中，在后续时间进行处理。通过评价器(evaluator)计算输入值的多项式函数来为近似曲线和曲面等几何物体提供有效的手段，然后由顶点描述的几何图元进行操作。在此阶段，对顶点进行转换、光照，并把图元剪切到观察体(frustum)中，为下一步光栅化(rasterization)作准备。光栅化产生的一系列图像的帧缓存地址和图元的二维描述值，其生成结果为基片(fragment)，每个基片适合在最后改变帧存之前进行操作。这些操作包括根据先前存贮的深度值进行有条件的更新帧缓存，进行各种测试以及融合。

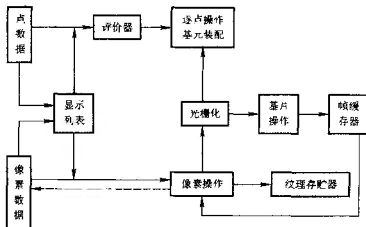


图7-7 OpenGL绘制原理

图像像素(像元)的处理包括像素、位图、影像等，它们经过像素操作之后直接进入光栅化阶段。

OpenGL的几何图元是由顶点描述的，以便于逐点操作并按其数据装配成基元，经光栅化形成基片。对于像素数据，其结果还可以存贮在纹理用的内存中，以便从中提取像素信息进行光栅化。OpenGL显示列表可以当作一组函数，它们存贮后能在后续时间内处理。OpenGL评价器也是许多特殊的函数，它们允许用户采用一个或两个变量的多项式映射来产生顶点坐标、法线坐标、纹理坐标及颜色，生成结果传递给执行管道(pipeline)。

光栅化包含几何和物理映射两部分。几何操作是将图元转化成二维图像；物理操作是计算图像每个点的颜色、深度等信息。因而光栅化一个图元由两部

分操作组成。第一部分是决定窗口坐标(window coordinates)中一个整数栅格的哪些方块由图元占有;第二部分是为每个这样的方块计算它们的颜色值、深度值。计算结果将传送到OpenGL的下一个过程,并用此信息更新帧存中的适当单元。

3. OpenGL 函数

Wndow95以及Windows NT3.51以上的操作系统提供了OpenGL动态库,VC ++ 2.0以上的版本提供了OpenGL静态库。在微机版本中,OpenGL提供了三个函数库——核心函数库(基本库)、实用库和辅助库。

OpenGL核心函数有100多个,都以gl开头。所有OpenGL提供的操作都可以使用这些函数来实现,而且对于不同的软件和硬件平台,这些函数的使用情况,包括语法和功能,都是相同的,这个特性注定了OpenGL完美的可移植性。

OpenGL的实用库是OpenGL基本库的一套子程序,它提供了40多个函数,这些函数都是以"glu"为前缀。同核心库一样,实用库也可以运行于任何OpenGL平台。

OpenGL辅助库有31个函数,都是以"aux"开头,它们不能在所有的平台上使用,但Windows95/98、WindowsNT支持它们。由于这些函数几乎不支持调色盘,没有为其他所需的消息添加句柄,而且消息不是在代码中进行,所以在产品程序中尽量不要使用它们。

在OpenGL中,还有一些以"wgl"开头的Windows专用函数,用于连接OpenGL和Windows窗口系统的。它们管理着色描述表(render context)及显示列表、扩展功能、管理字体等。

此外,OpenGL还有6个用于处理像素格式及缓冲的Win32 API函数和4个OpenGL结构。OpenGL这些函数提供了功能强大的图元绘制命令,所有高级的目标绘制都通过这些命令来实现。

4. OpenGL 与三维地形虚拟现实

许多虚拟现实的图形系统很早就使用OpenGL,如Performer、OpenInventor、Vega等。虚拟现实的核心问题之一是视景的生成。视景的生成首先应生成视景的内部条件,即建立视景生成的过程和对象的几何模型。接着完成视景基本内容的表示,如图形学表示、图像表示和运动图像表示。然后考虑视景的高层内部表示,如三维物体的几何模型、物体的结构层次表示等,最后完成真实视景的生成。OpenGL正是一种支持视景生成的程序设计语言。

用OpenGL显示三维地形和图形叠置主要步骤如图7-8所示,首先将Grid的三维数据读入计算机内存,由OpenGL生成三角片(GL_TRIANGLES)或四边形片(GL_QUAD_STRIP),组成地表三维模型,并设计网状显示、灰度显示和影像纹理显示三种不同的显示方式。为增强用户的临场感,可将雾气、光照以及

图像的纹理映射融入场景布局，使场景更加逼真。影像纹理显示采用材质贴图方式，将航空正射影像、遥感图像或数字栅格地图经纹理映射(确定每个空间——像片平面、景物空间——地面、图像空间——屏幕三个空间之间的映射)和投影变换，粘贴到地形地物表面上。

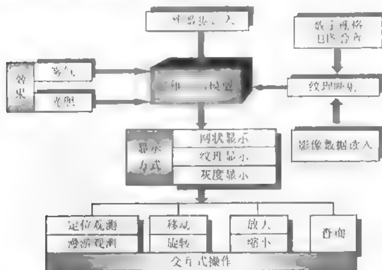


图 7-8 OpenGL 三维地形绘制原理

在虚拟现实应用日益广泛的今天，OpenGL 在虚拟现实中的应用地位得到公认，基于 OpenGL 的虚拟现实系统已经广泛应用于军事、航空航天、医学等领域等。可以肯定的说，OpenGL 正将人们带入一个三维世界。随着高新尖科技发展对图形、图像处理的实际需要，OpenGL 必将发挥越来越重要的作用。

7.2.4 VR GIS 地学应用实例

近几年，虚拟现实技术发展很快，通过与 GIS 技术结合，在地球科学领域得到了很好的应用，国内外有许多这方面的例子。例如，美国的 VR 技术专家和城市管理专家联合对洛杉矶和拉斯维加斯两个城市改造进行了虚拟城市规划实验，最后做出了人行道树种的选择。美国科学家运用 VR GIS 技术，将 DEM 与 1M 卫星影像进行叠加处理，对圣安得列斯断层进行了地震虚拟实验，逼真地显示了强烈地震的情景。美国全国超级计算机应用中心根据卫星遥感数据将大气、海洋与陆地的相互作用进行虚拟模拟，使之为天气预报服务。

浙江大学省资源与环境信息系统(GIS)重点实验室将地理信息系统技术

计算机数据库技术、互联网技术与虚拟现实技术等多种先进技术与流域管理规划管理相结合，自主开发了VR-GIS数字流域虚拟现实三维电子沙盘动态显示与分析系统平台，成功地用于水利流域规划、管理，建立了"黄河数字流域规划管理信息系统"。系统基于Client/Server模式，拥有自主开发的三维数据库引擎技术和VR-GIS技术，进行动态三维可视化开发。图7-9为数字流域三维电子沙盘的界面与场景，图7-10为三维可视化水库大坝规划设计的设计，图7-11为洪水淹没虚拟现实模拟，7-12为互联网环境下的虚拟现实场景的浏览，图7-13为城市虚拟现实建模可视化表达，图7-14为遥感影像与三维地表叠合显示。

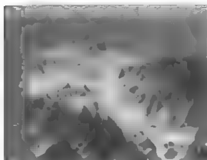


图7-9 数字流域三维电子沙盘

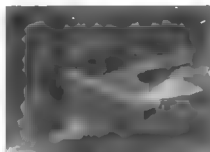


图7-10 三维可视化水库大坝规划设计

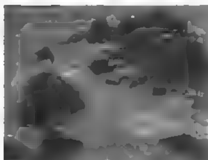


图7-11 洪水淹没虚拟现实模拟

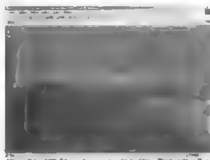


图7-12 互联网环境下虚拟现实浏览

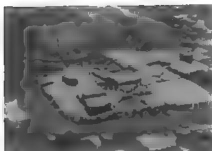


图7-13 城市虚拟现实建模可视化



图7-14 卫星影像与三维地表叠加

§7.3 时态地理信息系统

7.3.1 时态 GIS 基本概念

1. 地理空间数据和信息的时态特征

我们在§1.2.2.1讲过，地理空间数据和信息的三个基本特征，时态特征是其中之一。时态特征指地理数据采集或地理现象发生的时刻或时段，时态数据可以用来动态地表现地物的发展变化。按时间尺度，时态特征数据可以划分为短期(如地震、洪水、霜冻)、中期(如土地利用、作物估产)、长期(如城市化、水土流失)和超长期(如地壳变动、气候变化)等类型。不言而喻，空间数据时态特征的研究具有非常重要的意义。

不过，由于空间现象动态演变的研究难度，时态数据目前应用得还不多。换言之，现今绝大多数实用的地理信息系统一般仍是静态或准静态的，在这样的 GIS 系统中，时态数据通常作为普通属性数据的一部分处理。正因此，本书前面所有的内容都将时态数据作为属性数据的一部分对待，即一直未述及空间数据和信息的时态特征研究。

近年来，空间数据的时态特征越来越受到 GIS 学界的重视，有机又有地组织时间维和空间维的数据，实现 GIS 动态、多维化的问题提上越来越重要的议事日程。为此，我们觉得有必要在关于 GIS 新技术的第七章中，也介绍一下空间数据和信息时态特征研究方面的发展。本节要讲的时态地理信息系统，就是出于这种考虑。

2. 时态地理信息系统

时态地理信息系统，亦称为时态 GIS，Temporal GIS 或 TGIS。时态 GIS 是相

对静态 GIS 而言的，地理环境、事物和现象不断发展变化，但静态 GIS 仅对其进行“快照”式的表达。当研究的空间现象状态（如三维地形或其表面）发生变化时，静态 GIS 对原来的状态不予保留，只关心系统变化后的结果；它只是简单地对一定瞬间的地理现象进行备份，将发生的时间作为属性数据保留，而不惯性过去时刻状态的回放，也不对动态演变进行比较分析。这种只对当前数据状态进行“快照”式描述的 GIS 系统，称静态 GIS，或准静态的 GIS 系统。

时态地理信息系统将时间概念引入 GIS 中，跟踪和分析空间数据和随时间的变化。它不仅描述系统在某时刻、时段的状态，而且描述系统沿时间维变化的过程，预测未来时刻、时段系统将会呈现的状态，以此获得系统变化的趋势，或对过去不同时刻、时段的系统状态回放重现，总结出系统沿时间变化的规律。这种不同于静态 GIS，含有时间概念的 GIS 系统，称为时态 GIS 或 TGIS。

3. 时态 GIS 与时态数据库

时态 GIS 包含时态数据库（temporal database），内置有时空数据模型。时态数据库基于时态数据模型，是数据库系统技术的一个比较前沿的分支。时态 GIS 采用时态数据库技术，动态地研究三维地理空间的数据和信息。因此，时态 GIS 所基于的数据模型称为时态空间数据模型，或简称时空数据模型。

为此，我们必须首先介绍一下时态数据库的基本概念，特征或术语。

- 有效时间（valid time）。指一个对象在现实世界中发生并保持的那段时间，或者该对象在现实世界中为真的时间段。有时也称为对象时间、世界时间、外界时间和物理时间。

- 事务时间（transaction time）。指一个对象录入数据库系统的时间，有时也称为系统时间、数据库时间、历史记录时间、逻辑时间等。

- 用户定义时间（User-defined time）。指用户根据需要而输入的时间，数据库管理系统将它与库中其他一般数据同等对待。

- 快照数据库（Snapshot database）。使用删除、替换对数据库状态进行数据更新的一系列系统活动，它导致数据库的过去状态丢失并忘记，它仅反应现实中的一个片段。

- 回滚数据库（Rollback database）。根据事务时间在系统中保存对象的所有过去历史，并能对库中任何数据作更新操作。

- 历史数据库（Historical database）。根据有效时间在系统中保存对象的所有过去历史，并能对库中任何数据作更新操作。

- 时态数据库（Temporal DB）。根据有效时间和事务时间在系统中分别保存对象历史和数据库状态

- 元组时刻标记（instant-stamping of tuples）。每个元组含有一个时间值分量，时间值指明该元组为当前元组的开始时刻。这里，“元组”相当于“记录”

(元组及其分量的概念见§3.1.3.2)。

- 元组时区标记(interval-stamping of tuples) 每个元组含有两个时间值，这两个时间值分别指明该元组为当前元组的开始和结束时间。
- 元组时态元素标记(temporal-element-stamping of tuples)。每个元组含有一个时态元素，时态元素指明该元组的真的时间。
- 属性时刻标记(Instant stamping of attributes)。元组的每个属性值附加一个时间值，它指明该属性值的真的开始时刻。
- 属性时区标记(Interval stamping of attributes) 元组的每个属性值附加两个时间值，这两个时间值分别指明该属性值的真的开始和结束时间。
- 属性时态元素标记(Temporal-element-stamping of attributes)。元组的每个属性值附加一个时态元素，时态元素指明该属性的真的时间。
- 时空数据库版本。时空数据库的一个元组的有效时间区段，通过版本间的时空关系数据组织结构、查询历史数据。

4. 几种有代表性的时空数据模型

- 时空立方体模型。时空立方体模型用二维图形沿时间第三维发展变化的过程表达现实世界平面位置随时间的演变；给定一个时间位置值，就可以从三维立方体中获得相应截面、立方体的状态。也可扩展表达三维空间沿时间变化的过程。时空立方体模型的缺点是随着数据量的增大，对立方体的操作会变得越来越复杂，以致于最终对时空数据的处理速度变得很慢。
- 连续快照模型。此模型在快照数据库中仅记录当前数据状态，数据更新后，旧数据的变化值不再保留，即"忘记"了"现在"与"过去"的状态差别。连续快照模型是将一系列时间片段快照保存起来，反映整个空间特征的状态，根据需要对指定时间片段的现实片段进行播放。连续快照模型中的时态与空间数据间的关系简单，表达方便，在TGIS系统中易于实现。但是，由于快照将未发生变化的所有特征重复进行存贮，会产生大量的数据冗余，当应用模型变化频繁且数据量较大时，系统效率急剧下降。此外，连续快照模型不表达单一的时空对象，较难处理时空对象间的时态关系。
- 基态修正模型。为了避免连续快照模型将每张未发生变化部分的快照特征重复进行记录，基态修正模型按事先设定的时间间隔采样，只贮存某个时间的数据状态(称基态)和相对于基态的变化量。基态修正模型的优点是每个对象只需贮存一次，每变化一次，只有很小的数据量需记录；同时，只有在事件发生或对象发生变化时才存入系统中，时态分辨率刻度值与事件发生的时刻完全对应。但基态修正模型较难处理给定时刻的时空对象间的空间关系，对于将整个地理区域作为处理对象时，该模型处理方法难度较大，效率较低，管理索引变化很困难。

• 时空复合模型 时空复合模型将空间分隔成具有相同时空过程的最大的公共时空单元，每次时空对象的变化都将在整个空间内产生一个新的对象。对象把在整个空间内的变化部分作为它的空间属性，变化部分的历史作为它的时态属性。时空单元中的时空过程可用关系表来表达，若时空单元分裂时，用新增的元组来反映新增的空间单元，时空过程每变化一次，采用关系表中新增一列的时段来表达，从而达到用静态的属性表表达动态的时空变化过程的目的。时空复合模型的缺点是数据库中的对象标识符的修改比较复杂，涉及的关系链层次很多，必须对标识符逐一进行回退修改。

• 第一范式（1NF）关系时空数据模型。按第一范式（参见§3.1.3.2和3）方法，一个对象的历史过程可以通过几个元组来表达，元组中每个属性值都具有时间标记。第一范式关系时空数据模型的缺点是，在一个TGIS系统中，对于一个空间单元的表达，即使未发生空间拓扑变化，而仅一个属性特征值发生变化，就必须增加一个新的元组来表示，数据表中记录了大量重复数据。

• 非第一范式（N1NF）关系时空数据模型。这是相对于第一范式数据库方法提出的，处理非表格化复杂结构对象的时空数据模型。非第一范式关系时空数据模型的优点是，元组可以采用不定长和嵌套方式，对于复杂的空间单元的变化或整个演变历史，只需使用一个元组来模拟。所以，非第一范式关系时空数据模型非常适合时态GIS的应用。但是，由于时态数据库一些理论问题（如关系结构时态代数操作封闭性等）尚未很好解决，同时又受到可供实际应用开发的商业数据库软件的制约，目前很难采用非第一范式方法时空数据模型开发出实际运行的应用系统。

在上述各种时空数据模型中，目前比较具有实用意义的是基态修正模型。

7.3.2 基态修正扩展模型

基态修正模型不存贮每个对象不同时段的所有信息，只记录一个数据基态和相对于基态的变化值，提高了时态分辨率，减少了数据冗余量。在基态修正模型中，以矢量图形或栅格影像的处理模式进行时态属性查询时，难以建立直接索引关系。而基于时间、以事件驱动的面向对象的基态修正模型，与面向对象的关系型数据库能较好地结合，时间表与对象表易于关联，对于给定时间表中的元素，可即时查询出符合条件的基态变化量值。

1. 现时库、过程库和历史库

基态修正模型需要扩展。基态修正模型与其他几种模型存在一个共同的缺点，即它们关注的都只是给定时刻（时段）的数据状态，而不涉及由一种状态转变为另一种状态的过程。而且在实际应用中往往存在一种“伪变化”，也就是处理过程中经证实需要取消或修正的“变化”。这时，“伪变化”应被撤消，系统

需恢复、回退到变化前的状态。因此，需要设计"过程时态库"来记录、表达变化过程。通过在时空数据库中创建三个表达不同时态的数据库，即现时库（现在时态）、过程库（进行时态）和历史库（过去时态），是一条解决问题的较好途径。

下面就扩展的基态修正模型的现时库、过程库和历史库分别加以阐述。

• 现时库。根据基态修正模型的特点，处于最频繁操作的状态应作为基态，显然现时库中的对象一般应作为基态。在时空数据库中，现时库反映的是操作对象的现在时态的空间位置和属性，数据库中的每个元组都处在"激活"状态，是数据库的当前操作对象。

• 过程库。大多数的时态数据库中，只保存对象的空间位置和属性的现状（最后一次更新的数据状态）和变化前的各时段历史状态。然而合理的现实情况是数据的变化往往需要经历一个过程，必须满足一定的条件，经过过程操作验证后，才允许进入过程演变的下一个阶段，过程演变存入过程库中。过程库跟踪对象演变的所有阶段，描述事件发生和演变的全过程。同时，一旦事件发生的条件不满足时，该事件将会沿发生的时间轴回退，直到条件成立时停止或返回到事件发生前的状态。

• 历史库。一个对象当由一事件引起变化后，其最新状态被存入现时库中，它的演变过程被存入过程库中。当变化的所有条件成立时，事件将会按时间顺序排列，将相对基态修正量存入历史库中。为了对历史进行快速查询，对所有元组建立分级索引。对于任意给定的时刻或时段，都可从历史库中查询出"过去"的状态，恢复当时对象的时空关系，进行相应的时空运算。

2. 区段快速索引

根据时态数据库最新研究成果"时态数据分段存贮策略"，在基态修正区段索引中，时态数据库用户对数据具有厚今薄古或对某一特定时期操作频繁的特点。根据这一客观事实，将时态元素的历史分为远、中、近三个时段，即"三史"制，如图 7－15 所示。根据时态应用领域管理的特殊性，常常将三史制扩展为多史制。

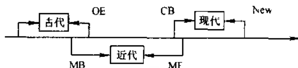


图 7－15 时间三史制区间表达

提高时态数据库查询效率的一个较好的解决方法是建立区段快速索引。建立多区段的快速索引的原则取决于基态修正模型的分辨率刻度值大小、基态

修正的频率、事务处理时间存贮制式(现代、近代、古代,历史制)的跨度,以及数据库的总体规模等因素。图7-16说明基态修正区段快速索引的基本概念。Δfi为两相邻时态元素对象的相对基态修正值,即"差文件"deltafile的值,Δfci为区段相对基态修正值。时态元素对象、基态修正区段快速索引,以及区段相对基态修正值通过时空数据表发生关系,并相互作用。利用时空数据库中建立的区段快速索引表,可快速将指针指向相应时空元素表中的,具有公共区段相对基态修正值属性的对象。图7-17表示索引机制与时空数据文件的关系,采用由Δfc建立起来的快速索引标记,大大提高了数据库查询速度,成倍减少了图形数据的计算量。

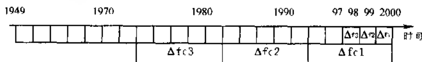


图7-16 基态修正区段索引

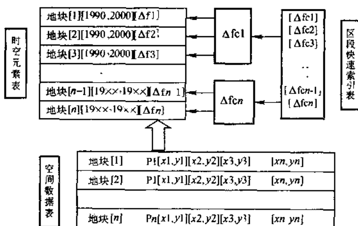


图7-17 区段快速索引与时空数据文件相互关系

3. 变粒度存贮因子

为了提高查询速度,在基态修正区段索引中,引入变粒度索引因子。根据时态元素的变化频度,确定相应的变粒度索引因子K值,如图7-18所示,选择最佳的区段起点和终点,K值越大,相对基态索引标记越密集,本区段内检索速度也就越快。由此获得各基态修正区段长度,并建立相应的索引表。系统运行表明,通过变粒度索引因子的引入,系统查询效率能够明显提高。

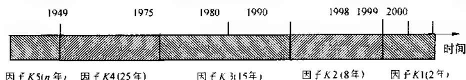


图7-18 变粒度存贮因子与时间区段

7.3.3 基态修正扩展模型的应用实例

土地产权产籍管理信息系统是时态GIS的最典型的应用系统之一，下面根据土地管理时空变化特点，详细介绍基态修正扩展模型在系统研究开发中的应用方法。

1. 土地时空过程分析

土地地籍管理信息系统所管理的对象涉及到土地空间位置和权属以及权属变化的情况。土地权属变化是由土地的划拨、产权变更、产权转移等情况引起，这种变化涉及到时间、空间和属性的变化，归纳起来有三种情况，一是地块的空间形状、状态和性质同时变化；二是地块的空间形状、状态发生变化，但性质不变；三是地块的状态或性质发生变化，但空间形状不变。土地在划拨、征收和交易过程中，其空间变化的基本形式是“分割”和“合并”。地块分割是指一块地由于权属或用途的变化被分成两块或多块的情况；地块合并是指两块或两块以上地块由于权属或用途的变化被合并成一块的情况。随着时间的推移，土地分割与合并的组合构成复杂的历史变化序列。

图7-19表示一假想街道地块分合情况和各时段权属情况。地块的时空变化是非常复杂的，然而引起这些变化的主要因素是人（权利人）和人的行为（征收、划拨和交易）。土地管理信息系统对时空数据的管理是以产权变化为时段标尺，通过记录和规范人（政府土地管理员）的工作行为（工作流程），实现对产权属性数据和空间图形数据的统一管理。

无论是空间实体还是属性实体都存在一定的生命周期，宗地的生命期定义为某一权利人持有产权证的时间。各实体的变化和时间是紧密相关的，因此各实体内部还存在一种时态（temporal）关系，系统必须能有效的管理各实体以及实体与实体之间的动态时间关系。通过对现有时空数据模型的深人研究，作者认为，在城镇土地管理系统中采用基态修正模型较为适宜，同时对其作相应的改进，以提高时空数据查询的效率。

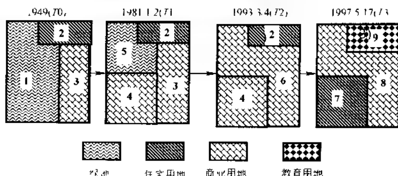


图7-19 地块变迁示意图

2. 土地时空数据结构

时态空间数据和属性数据有机地结合是建立土地信息系统时空数据模型的关键。土地管理信息系统，其主要特点是数据流动过程中所涉及到的地物实体的空间信息和属性信息的紧密结合，充分利用现有地籍调查成果，为系统数据库的动态更新提供丰富的信息源。

根据土地管理信息系统的系统功能要求，需要建立各类与之相适应的数据库。地理空间定位基础数据库以空间数据为主，结合少量的属性数据，便于一般查询及作为参照背景，如地形图数据库、控制网库等。业务事件驱动监控数据库则是纯属性数据库，无空间数据的查询。宗地数据库是集空间数据与属性数据于一体的数据库。根据宗地、宗地属性及其相应的时态数据库特征，需采用一种全新的数据模型进行描述，即具有时态特征的空间数据和属性数据模型，以满足系统的各种功能，这种数据模型能够表达时空混合数据关系。图7-20为宗地时空变化的基本类型。

土地时空数据结构包括下述主要成分：

- 宗地空间数据。宗地在空间表示为位于地面的面状物（多边形），其空间位置可由其边线的一系列大地坐标串来确定，每块宗地用地均有其宗地编号。因此，空间数据可以表示为：宗地编号1，边线1，(X1，Y1)，(X2，Y2)，(X3，Y3)…；边线2，(X1，Y1)，(X2，Y2)，(X3，Y3)，…；宗地编号2，…，等。
- 宗地属性数据。宗地的属性数据主要包括宗地权属、土地用途、工程项目名称、容积率、绿化率、建筑密度、用地面积、用地性质等控制指标。这些属性对应于不同的建设用地各不相同，可表示为：宗地编号、权利人名称、用地性质，用地面积，…这些属性与空间坐标 X、Y 无直接关系。
- 宗地时空、时属关系。随着时间的推移，必然有征地、划拨、交易等事件

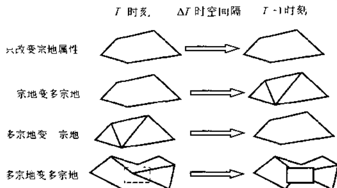


图7-20 宗地时空变化基本类型

不断发生,这些事件可引起宗地实体的空间或属性的变化。

根据以上两种不同类型数据的分析,要实现对宗地数据和公共专题数据等的多用途数据的查询显示及相关的空间分析和操作,需建立一个这样的数据模型,其满足如下的条件:

其一,存贮空间数据,能输入、输出、编辑;

其二,存贮属性数据,能输入、输出、编辑;

其三,存贮时态数据,能输入、输出、查询、带限制条件的编辑;

其四,建立空间数据、属性数据和时态数据的交叉双向链接。

● 时空、属性与核心关联表相互操作。将空间数据上键字段、属性数据上键字段和时态数据上键字段组成一个核心关联表,无论仅空间数据发生变化,或仅属性数据发生变化,或空间属性均发生变化,将都视为时空数据发生变化,相应的数据表(一个或多个)变化量被记录,核心关联新表生成,变化前核心关联表被压入历史数据库。通过指定时态变化值或某一确定时段,就可以通过空间数据来查询或修改属性数据,也可以通过属性数据来查询或修改空间数据。若指定的时态时段为过去时态,则可根据历史数据库中的核心关联表,恢复当时的时空关系,并实现由空间数据查询属性数据,或通过属性数据查询空间数据,但不能修改历史数据。图7-21为时态、空间及属性数据与核心关联表的相互操作关系。

空间数据除了宗地数据外,还有道路、水系、控制点等线、点状数据,这些数据的空间查询和分析也是必不可少的。系统数据库数据模型必须包含点、线、面等特征的空间数据,并建立明确的定义和合理的空间拓扑关系,数据的存贮必须和现行的数据库软件兼容并且能进行相互数据访问。

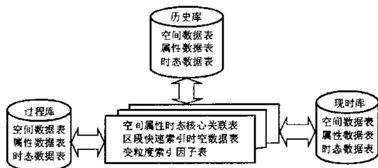


图7 21 时态、空间及属性数据与核心关联表的相互关系

3. 区段快速索引及变粒度索引因子的实现

在历史库和现时库的空间数据表、属性数据表及时态数据表中，存贮着时间标记。区段快速索引时空数据表 $\Delta fc(i)$ 值及变粒度索引因子表 $K(i)$ 值通过数据字典导人设定数据，对历史库、现时库中的带有区段标记和变粒度因子属性的元素记录依据 $\Delta fc(i)$ 和 $K(i)$ 值进行快速检索，而不是通过基态文件及逐个差文件(deltafile)进行运算获得查询结果。

4. 数据模型的拓扑关系及时空数据存贮

空间数据可以矢量和栅格两种方式存贮，根据地籍管理工作和数据的特点，需采用矢量格式存贮空间图形数据。任意的平面图形都可以由点、线、面三要素构成。因此，土地系统数据也分点、线、面来描述空间数据，并建立如下拓扑关系：

- 点。用户标识码(USER - ID)以及点的 X、Y 坐标表示。
- 线。用户标识码和线上的一系列 X、Y 坐标表示。
- 多边形。用户标识码和其周边的线系列表示。

不同专题内容的空间数据分层存贮，即一类专题层为一类空间数据，如建筑物层，道路层、控制网层、宗地层等。不同专题内容的空间数据可通过统一的空间坐标系进行定位，即采用一组公共的控制点或同名地物点，进行最小二乘法平差计算，实现不同图层的精确透明叠加。

系统可采用现行成熟的大型关系型数据库进行时空过程管理，如 ORACLE、SQL Server，大型关系型数据库都具有较强的网络能力和数据库管理能力。对于系统数据库中所采用的数据模型，空间数据和属性数据之间能否建立合理的连接关系，对能否实现两者之间的双向查询和分析至关重要。源于关系数据库多表连接机制的启发，在数据模型中建立时空数据和属性数据之间的关系机制，系统数

据库空间数据的数据组织设计为 ID,X,Y…的形式。为此建立一个时空与属性关系表,包括空间数据的 ID 和属性数据表中的关键字,通过这个表实现时空与属性双向查询、分析等操作。空间数据与属性、时态数据共存于关系数据库中,关系数据库对图形数据的存贮管理机制可建立如图 7-22 所示的时空关系。

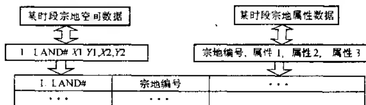


图 7-22　宗地时空、属性关系

5. 土地产权产籍管理信息系统应用实例

根据前述的时空数据存贮方法,浙江大学省资源与环境信息系统(GIS)重点实验室开发了一套大型网络级动态土地产权产籍管理信息系统(简称 ReGIS)。系统由 10 类数据库和 10 余项主要功能模块组成,涵盖了土地产权产籍管理的全部业务,如图 7-23 所示。

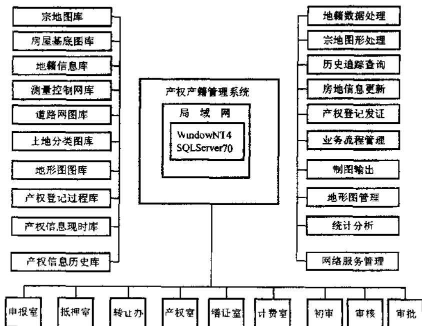


图 7-23　土地产权产籍管理信息系统 ReGIS 系统构成图

该系统基于Oracle/SQL Server数据库平台，采用VC++6.0开发语言，及COM控件技术，内置时空数据模型，即现时库、过程库、历史库，兼容多种GIS平台数据格式，具有强大的GIS图形编辑功能，动态管理土地管理工作的全部业务流程，方便地查询现状、过程、历史图形属性拓扑信息，实现了对土地管理工作登记、调查、审批、增止全过程的动态管理。系统可任意恢复给定时刻的历史时空关系，查询统计并输出各类土地信息，具有很强的数据安全备份机制，自动定时备份并能恢复受破坏的数据库。用户可方便地修改已十分丰富的数据字典，超级用户可根据不同的业务流程和科室职能对系统功能进行组合调整。通过系统提供的电子跟踪卡功能，领导及业务人员可方便地对每宗业务进行动态实时跟踪管理，有利于领导督办业务及监督办事过程。系统具有WebGIS功能，可通过互联网发布土地信息、发送土地数据，用户可在网上查询所需要的土地信息，土地局可通过网上受理土地交易业务，实施分布式网络管理。图7-24为系统的后台服务管理器现时库、过程库、历史库数据维护模块。图7-25为从历史库中查阅土地产权的历史变更信息，界面显示的是当时办理土地转让业务的调查资料。

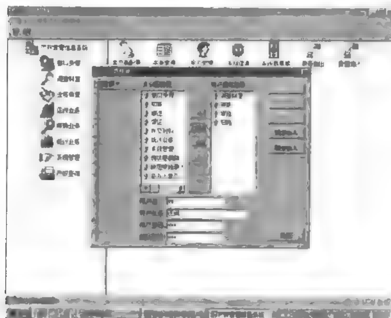


图7-24 现时库、过程库、历史库数据维护

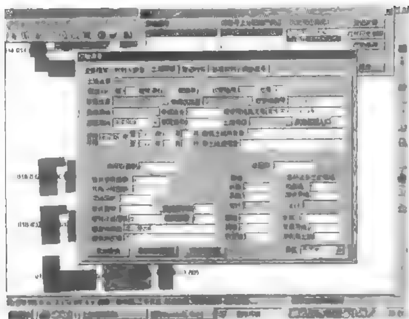


图7-25 历史库中查询变更信息

§7.4 客户/服务器体系结构及空间数据库引擎

客户/服务器(client/server,C/S)体系结构和空间数据库引擎是目前新一代GIS应用系统的核心技术之一。前者确定了系统的网络结构模式和数据库类型,后者使GIS专业应用系统和具有客户/服务器结构的通用商业数据库系统进行无缝连接。无论是系统的稳定性、安全性、开放性、一致性、可靠性、开发和可扩展性,还是系统的友好性和总体功能,与过去的应用系统相比较,都有了很大的改进和提高。

7.4.1 客户/服务器体系结构

近年来,在GIS领域利用客户/服务器结构开发的应用系统已经逐渐普及,代替了传统的基于主机或文件服务器的处理结构。客户/服务器结构以其自身的优势已经被越来越多的开发人员和应用单位所采用。

1. 客户/服务器基本概念

在早期的计算机网络结构中，人们主要采用集中式处理结构和文件服务器结构。集中式处理结构的核心处理均在处理中心的主计算机上进行；集中式结构也可以使用 Oracle 或 SQL Server 等数据库系统，但应用程序和数据库位于同一主机进程中。而在文件服务器结构中，网络的服务器端唯一所需的，是一个简单的文件服务器，所需的服务仅仅是文件服务。数据库在文件服务器网络中的运行就像在本地硬盘上运行时一样，服务器端没有数据库引擎；数据库引擎实际上运行在客户机上，与客户应用程序紧密联结。当用户要求引擎执行 SQL 操作时，引擎使用标准的文件读写操作命令，在网络上搜索数据库文件；整个表在网络上传递，搜索工作在客户端最后完成。

客户/服务器结构是一个活动的数据库服务器进程。在这个数据库服务器中，运行着如 SQL Server、Oracle 等数据库系统。在客户/服务器结构中，客户端向服务器传送数据请求，然后服务器处理该数据请求，最后只把结果返回到客户端，如图 7－26 所示。

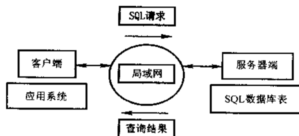


图 7－26　客户/服务器模式运行模式

GIS 应用系统由过去的基于文件服务器方式转变为基于客户/服务器方式，除了能享受客户/服务器方式本身带来稳定、安全性、可靠、开放、并发等好处外，还具有下述三方面的优势。

其一，当数据库服务器开始处理查询操作时，响应时间由服务器计算机主机决定，而与客户机无关。系统可以配置一台高档服务器，供所有客户端用户分享。

其二，活动的服务器进程只把结果集合，通过网络传回，而不是大的整个文件，因此网络负担大大减轻，效率显著提高。

第三，活动的服务器进程能更好地保证数据的一致性。文件服务器数据库在有限的范围内，也能通过客户应用程序数据库引擎访问基本的原始数据。但只有在客户/服务器结构数据库中，强大的服务器引擎才能够访问所有原始

数据。

2. 典型两级客户/服务器开发模式及其局限性

近年来在一些的中小型 GIS 信息系统开发模式中，有不少采用第四代开发语言(4GL)，如 VC、VB、PB 等开发的客户/服务器应用系统。它们在逻辑上可划分为两层：前端和后端，如图 7-27 所示。

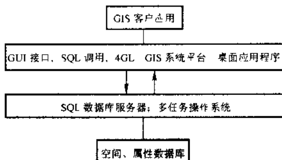


图 7-27 典型的两级客户/服务器模式

前端进程使用 Visual Basic、Visual C ++、Power Builder 等第四代开发语言开发。在前端用户只与 PC 机进行交互。前端的可执行代码由 GUI 图形用户界面元素组成，如按钮、菜单、GUI 窗体流、SQL 语句、数据验证等。

后端进程由 SQL Server 或 Oracle 数据库服务器多任务操作系统组成。后端数据库的内容包括数据表、触发器、引用一致性定义、安全策略、备份策略等。

在两级客户/服务器系统中，逻辑前端展开在基于 Windows 的客户工作站上，逻辑后端展开在服务器计算机上，逻辑结构与物理结构一对一的映射。正是由于这种一对一的映射关系，产生了以下几个主要方面的局限性。

第一，难于管理前端的增强功能。在两级环境中，如果在数据库中新添一列(如宗地信息表)，就需要修改 SQL 查询，即随着数据库的改变，所有客户工作站需要同时升级为新版本前端。

第二，难于共享公用过程。当根据用户应用需求设计完商业过程，如土地产权产籍系统中的存贮过程，并根据其流程规则开发完应用程序后，很难在另一应用中重新使用这个过程。也就是说，两个完全相同的商业过程必须在两个不同的应用程序中维护。

第三，难于保证安全性。在两级环境中，安全性主要由数据库服务器后端的安全性和客户应用程序前端的安全性来保证。数据库安全性主要是把权限分配给数据库对象的用户，有的用户只有 SELECT 权限，有的用户有 SELECT、UPDATE 或 INSERT 权限。对于一个复杂的、有许多用户和部门的系统而言，尤

其是有些用户同时具有多重不同权限的角色，确定权限分配不仅困难，而且极易出错。

在应用程序的保护安全性中，流行的方案是使用数据库用户 ID 和口令登录到数据库中。但由于没有保护数据库的对象，应用程序安全性的实现必须通过系统中所有自定义应用程序开发和维护。

第四，无商业服务处理的数据库引擎。在两级客户/服务器结构中，前端和后端无法让应用程序开发人员方便地开发保持商业规则灵活性的后端过程，没有在多个数据库环境中管理事务的监视程序。

3. 三级客户/服务器服务模式

为了克服二级客户/服务器模式存在的局限性，数据库厂商及系统开发商都开始转向三级应用程序结构。新增加的一级位于前端和后端之间，为商业（泛指所服务的业务）规则提供了明确的中间层次，封装了与系统关联的商业模型（商业规则），并与用户应用层和数据库代码分开。三级开发模式中各级服务的作用如表 7-4 所示。

表 7-4 三级开发模式的各级服务

级	服务类型	特　性	任　务	工　具
一级：用户服务	客户应用程序	GUI 界面	表示和定位	4GL 开发工具及图形平台
二级：商业服务	商业服务器	商业对象属性和方法	商业策略、规则和安全性	4GL、VC、VB COBOL 等
三级：数据服务	数据服务器	原始数据管理	与决策无关的数据完整性	数据库或消息发送系统

在三级服务模型中，任一级可在同级或上下级间通信，上一级向下一级发出请求，其结果可返回到同一级内的进程，也可以返回到上一级或下一级的进程中。客户应用程序不能直接同数据服务通信。三级客户/服务器中各级的作用分述如下。

- 客户应用程序。客户应用程序包括：用户服务、客户、应用程序、前端和 GUI。客户应用程序的功能是为用户提供商业服务的图形界面，将商业服务作为一个整体理解，并高效地定位所需的商业服务。

- 商业服务。商业服务提供客户应用程序与数据服务之间的联系，其作用是执行商业策略，封装商业模型，并把该模型表述给客户应用程序，维护第三级中数据服务的一致性，把客户应用程序从复杂的原始数据服务中分离出来。通过在商业服务器级执行商业策略，系统就不必再在数据服务器级中执行商业策

略。在两级结构中，商业策略必须同时由前端的应用代码和后端的数据库对象来执行。

- 数据服务。第三级由数据服务器组成，最常见的数据服务器如基于SQL Server、Oracle等的数据库管理系统。数据服务器管理原始商业信息，并为商业服务器处理事务，保证数据的完整性。

下面给出一个采用三级客户/服务器模式开发的土地产权产籍GIS系统的实例。该系统的三级结构如图7-28所示。

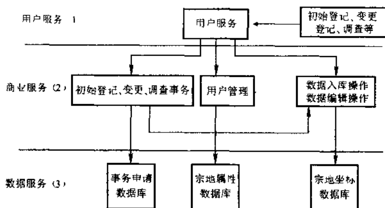


图7-28 三级模式下的土地产权产籍系统

7.4.2 空间数据库引擎

1. 空间数据库引擎的概念和作用

空间数据库引擎（spatial database engine，简称SDE），这一用语出于ESRI公司，指GIS系统借以实现基于大型关系数据库的客户/服务器模式的软件，在一些GIS文献中亦称为中间件软件。

我们在§3.2.2讲过，传统GIS采用关系数据库与文件系统混合管理的模式，将空间图形数据与属性数据分别加以组织和管理：用关系数据库直接管理属性数据，用文件系统组织图形数据和管理图形库；两种数据通过唯一标识符进行链接。采用关系数据库与文件系统混合管理模式的传统GIS数据库系统技术，虽然有重要意义并取得应用上的较大成功，但由于不得不部分地采取较落后的文件方式管理数据，总体上还达不到通用数据库技术在冗余度、数据独立性、共享性和易扩展性方面的要求，更不能享受大型关系数据库进步所带来开放性、一致性、安全性、标准化等一系列重要优势，不适合管理海量、多用户的

网络级的数据库系统(见§3.1.2.2)。

为此,GIS 专家一直在寻求更好的方式来管理空间数据。但是,由于空间数据的复杂性,用关系数据库来管理空间数据有一定困难(§3.2.1)。事实上,基于大型关系数据库的客户/服务器技术早在 20 世纪 80 年代末期就已基本成熟,成为商业信息系统运行中的核心功能。由于功能强大,性能稳定,并具有安全性、完整性、多用户共享机制和完备的工业标准,像 Oracle、SQL Server、SYBASE 和 Informix 等这样一些企业级关系数据库的 client/server 技术,从 80 年代末期就开始在企业级大型信息系统中广泛应用。但是,将空间图形数据存入关系型数据库中,充分利用关系型数据库的完备的运行机制,实现图形、属性数据后台统一管理的问题,直到 90 年代中后期才开始解决。

解决问题的途径就是开发一种中间件软件,它的作用就是为空间图形数据享用大型关系数据库架起一道桥梁。1996 年后,GIS 基础软件开发商和大型数据库开发商在联手的基础上相继推出了这方面的产品,并借用"数据库引擎"的基本概念,取名空间数据库引擎(spatial database engine,简称 SDE)。目前,国际上比较著名的 SDE 类产品有:GIS 软件开发商 ESRI 公司的 Spatial Database Engine (SDE)、Oracle /MapInfo 公司联合开发的 Oracle Spatial(两年前为两个不同的产品:MapInfo 公司的 SpatialWare 和 ORACLE 公司的 Spatial Cartridge)以及 Intergraph 公司的 GDO 产品。

实现将地图数据存入数据库这一目标,至少先解决三个问题。

第一,要解决在关系数据库中存贮空间数据的问题。空间图形数据是非结构化的复杂的,要利用规范化的关系数据库来存贮空间数据,需要进行特殊的努力。

第二,将数据存贮到数据库中仅仅是一小步,还必须实现对数据库服务器中的空间数据的空间查询、检索功能。

第三,还要提高空间查询和检索的效率。为此,必须建立有效的空间索引。

通过关系数据库存贮并有效检索空间数据,可能采取两种途径,但每种途径都会有不少的技术问题需要解决(参见§3.2.4.1 小字)。

总之,只有实现了上述功能,才能说是形成了一个真正的空间数据库服务器系统。将地图存入数据库不仅是一种存贮方式的改变,更标志着地图已经从一种非核心数据发展到一种核心数据。在数据库服务器中,其他数据类型与空间数据类型相结合,能更有效地利用数据、分析数据、发掘出隐藏在数据中的信息。

下面主要介绍 ESRI 公司的 Spatial Database Engine (SDE) 和 Oracle /MapInfo 公司的 Oracle Spatial 平台。

2. ESRI 公司的 ArcSDE

ESRI 公司在其代表性的不同版本地理信息系统平台 Arc/Info 的基础上，推出了针对国际上各主要大型数据库平台的空间数据库引擎 ArcSDE（由于系统不断升级，产品名称不断变化）。它提供了应用程序接口（API），使得对空间、非空间数据进行高效率操作成为可能。在客户端，用户可以通过 C－API、ArcView、MapObjects 等开发客户端应用程序。在服务器端，有 ArcSDE 空间数据引擎（应用服务器）、RDBMS 的 SQL 引擎及其数据库存贮管理系统。ArcSDE 通过 SQL 引擎执行空间数据的搜索，将满足空间和属性搜索条件的数据在服务器端缓冲存放并发回到客户端。

ArcSDE 将数据库中的各层图形数据划分为不同单元，建立相互对应的索引表，并对每个单元进行描述，实现对图形要素在数据库中的定位。具体的做法是：将空间数据类型加到关系数据库中，但不改变数据库平台的结构，只是在现有数据表中加入图形数据项（shape column），真正的地理数据放在另外的表中，通过关键字段与之关联，实现系统的访问和操作。ArcSDE 能处理的几何要素为离散点、简单线、复杂线和多边形面，可以在二维或三维空间上处理数据。ArcSDE 可以在 ORACLE、SYBASE、SQL Server、Informix 和 IBM DB Ⅱ 等数据库平台上运行。

ArcSDE 可通过 SQL 引擎提取数据子集，其速度仅取决于数据子集的大小，而与整个数据集大小无关。因此，ArcSDE 可以管理海量数据。另外 ArcSDE 还提供了一种不通过 ArcSDE 应用服务器，直接访问空间数据库的连接机制。这样，不需要在服务器端安装 ArcSDE 应用服务器，由客户端接口直接把空间请求转换成 SQL 命令发送到 RDBMS 上，并解释返回的数据。ArcSDE 在服务器和客户端之间数据传输采用异步缓冲机制：缓冲区收集一批数据，然后将整批数据发往客户端应用，而不是一次只发一条记录。在服务器端缓冲处理的方法大大提高了网络传输效率。ArcSDE 基本结构原理如图 7－29 所示。

3. Oracle /MapInfo 公司的 Oracle Spatial

Oracle Spatial 是 Oracle 公司和 MapInfo 公司共同开发的一个 Oracle 数据库内置模块，运行于数据库服务器上，采用空间数据库引擎技术，实现了高效的空间索引技术（R－tree），能够较好地处理海量地图数据。Oracle 8i Spatial 在 Oracle 8i 大型数据库环境下，基于 SQL 进行空间查询和分析的空间信息管理系统，不仅实现了在 Oracle 中存贮空间数据类型（点、线、区域等），而且建立了一套基于 SQL 的空间运算符，使空间查询和分析能在服务器端进行，其高效的基于 R－Tree 的空间索引技术则保证了空间查询的快速和准确。

Oracle Spatial 的系统结构原理如图 7－30 所示。系统主要由三个部分组成：MapInfo Professional（或 MapX）、Oracle 8i 数据库管理系统和 Oracle Calling

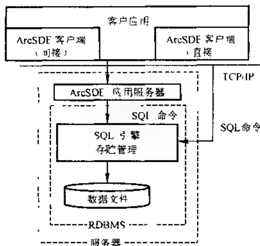


图 7-29　ArcSDE 基本结构原理

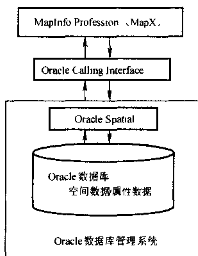


图 7-30　Oracle Spatial 的系统结构原理

Interface（OCI、Oracle 调用接口）。利用 Oracle Spatial 模块，用户可以将地图数据存贮到 Oracle 数据库中，在客户端通过 MapInfo Professional 或 MapX，通过 Oracle Spatial 提供的空间算子和 SQL 语言，对 MapInfo 地图进行编辑、查询和其他复杂的空间分析。Oracle Spatial 利用数据的内在空间关系，使得数据可视化和

空间分析系统汇入到关系数据库系统中。同时，Oracle Spatial 建立了 Client Server 结构的空间信息系统，不仅解决了海量数据的存贮和管理问题，同时解决了多用户编辑、数据完整性和数据安全机制等许多问题，给 MapInfo 的应用带来了极大方便。

在服务器端，Oracle 8i Spatial 将地图对象作为一个单独的列添加到 Oracle 数据库的表中，使得现有的 Oracle 数据地图化，其他没有地图的数据也可和数据库中的地图数据进行关系连接，以实现双向查询。MapInfo 公司的多数产品，包括 MapInfo Professional 和 MapX 都可以作为 Oracle 8i Spatial 的客户端程序，通过 OCI 对 Oracle 数据库进行访问，得到空间数据和属性数据，并加以表达。

Oracle 8i Spatial 的客户端可选用标准的 ODBC 接口，或 Oracle 调用接口访问 Oracle 8i Spatial 中的地图数据。采用 ODBC 和 OCI 是外部程序对 Oracle 数据库系统访问的两种常用方法。虽然 ODBC 在使用上比 OCI 要方便，但是从灵活性和访问效率的角度来看，OCI 的性能要比 ODBC 优越得多。

7.4.3 自主开发空间数据库引擎

虽然国外 GIS 平台软件商相继推出了功能较强的空间数据库引擎，但由于其价格非常昂贵，加之极其严密的版权保护措施，给合法用户带来了极大的不便。此外，每一种引擎本身或多或少带有软件商自身产品的局限性，所采用的数据格式基本上与自己的主流产品的数据格式一致。这样，系统的整个数据库建库及前台开发受到了制约。

在中国 GIS 产业的整个应用领域，占绝大多数的是中小规模的广大用户，他们往往对 GIS 应用需求非常迫切，但不愿意承受较大的软件开支。他们对 GIS 软件功能的要求往往比较集中在与自身业务关系较大的某一方面，对其他与业务无关的功能不关心。这样，有针对性地开发具有一定应用特点和功能、精简高效的空间数据库引擎，具有十分重要的现实意义。精简型空间数据库引擎不仅提高了整个 GIS 系统的实用性，实现图形和属性数据后台统一管理，使 GIS 系统上升到大型网络级客户/服务器模式的层次上，而且性能价格比高，市场竞争力较强，易于推广。同时，由于是自主开发产品，也为满足用户提出的特殊需求以及特殊的应用对象要求(如基于时态数据库的土地产权产籍管理系统)提供了很好的条件。

下面介绍一个自主开发的精简型空间数据库引擎实例。该引擎称为 Spatial Database Guide(SDG)，是浙江大学省资源与环境信息系统(GIS)重点实验室开发的，用于土地产权产籍管理系统的空间数据库引擎。

1. SDG 数据模型组织及操作

在关系型数据库中，对空间图形数据管理必须采用连续数据结构。SDG 的

结构模型，是将空间数据类型加到关系数据库中，以实现在关系数据表中对空间数据的操作，但不改变和影响数据库内核，仅在现有的数据库表中加入图形索引项，创建空间数据表，供系统调用及访问与其关联的空间数据。如图7-31所示，空间数据存入关系型数据库中，并实现查询的基本概念。

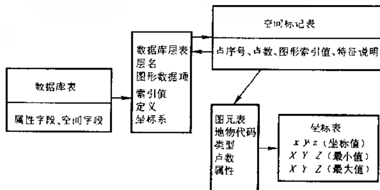


图7-31 SDG结构模型数据组织及操作

在图7-31中，层表(layer tables)将区域内的地理数据分成不同的专题层，每个专题层的具体数据可以存放在不同的表中，层表记录了区域内空间数据的基本状况，并对其他的表进行管理。通用的商业数据库表，加入图形数据项后，称为空间标记表(spatial stamping tables)。

空间标记表的每一条记录对应一个图元，除了记录若干属性值外，还包含图形的数据字段(figure)。但实际的空间数据并不放在空间标记表中，而是存放在相关的图元表(features table)中，两者通过figure字段关联。

图元表(features table)是记录空间数据的表，它对每一个图元进行编号，记录每一个图元的类型(点、线、多边形等)、该图元所含点的数量(对线而言是端点和节点、对多边形而言是顶点)和点序等。但具体的坐标值不放在图元表中，而放在坐标表中，两者通过图元号(FID)关联。

坐标表(coordinate tables)存放每一个图元的点坐标，以及图元最大、最小坐标值，它通过图元号与图元表关联。采用这种方式存贮和组织空间数据后，对空间标记表就可以像对待通常的关系表那样进行图形数据的合并、分割和拓扑运算，也可以进行图形到属性或属性到图形的查询。

2. 数据模型的空间图形要素表达

采用连续的数据结构，所管理的数据对象可以不受范围的限制。可将连续相连的封闭图形构成一个大的封闭图形，然后拼接成整个区域。在数据库中，

每一个封闭图形被作为一个完整的独立图形存放，只需一次操作，就能实现对整个图形的提取，具有相同属性的图形存放于同一个图形数据层中。数据模型可将整个层定义为可编辑、可显示，或不可编辑、不可显示等，所有的几何形状可归纳为点、线、面三种基本要素。

• 二维平面空间点。定义离散的、无面积或长度的地理要素，如大比例尺地图上的水井、电线杆以及小比例尺地图上的建筑物等，其拓扑关系可用单一的(x,y)坐标记录表达。

• 简单线段。一组有序的(x,y)坐标串表示的地理要素，如街道、河流、等高线等地理要素。

• 面(多边形)。一组起始点和终止点相同的封闭线段对应的(x,y)坐标串或称多边形，如行政边界、土地利用图地块、地籍宗地等。面可以是简单的多边形或带岛的多边形。

二维坐标值还可以扩展到三维。z值用来表示(x,y)点处对应的高度或深度；图形可以是二维的(x,y)，也可以是三维的(x,y,z)，对每一种类型的图形都有一组合法性检查规则，用以在将该图形存入 RDBMS 之前，检验其几何正确性。

3. 数据模型主要技术条件

数据模型实现的主要技术条件包括以下几个方面。

• 数据模型组织。建立一整套完善的空间数据模型体系，定义能满足空间查询和分析运算的地理数据类型，将空间数据合理地存入关系型数据表中，并通过双向索引表建立商业数据表与空间坐标表间链接，解决在一个 database 内多个属性与空间 table 的关联一致性和可逆性问题。

• 数据空间完整性操作。必须实现标准 SQL 操作下对数据的锁定机制，解决 Client/Server 结构下对数据的共同操作，防止用户间破坏数据完整性问题。

• 数据库中实现数据保护。必须确保将空间数据存入数据库中无需进行复杂的应用修改或使用特殊的保护装置，对用户使用权限的管理能够简单方便地实现。

• 优化数据库查询。使用基于规则的查询优化技术，实现对空间数据的有效操作，提高客户/服务器交互处理速度，达到或超过空间数据库平台的空间查询能力。

• 建立供开发必需的客户库和商业服务库。建立一套在客户端、商业服务层供开发工具调用的函数库，即 C - API，建立空间数据表和属性表的合并关系，支持空间数据表处理，即当系统访问空间数据表时，使用系统提供的 API，访问属性数据表时使用标准的 SQL 功能。

4. 数据模型体系结构

SDG 采用三级 Client/Server 网络系统结构，数据模型在客户应用端、商业服务端和数据服务器端均能对空间图形数据进行操作。其体系结构如图 7－32 所示。

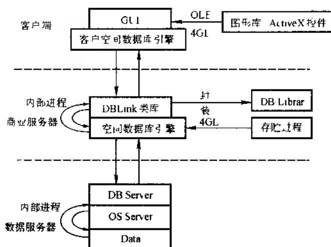


图7－32 数据模型的体系结构

在客户端可以利用 Visual C ++ 调用 ActiveX 控件，或为某一特别应用直接进行开发，应用程序调用的客户空间数据库引擎应用库是一组功能强大的类库，是一个处理客户端应用请求的 API 接口。

商业服务接收应用服务发出的请求，通过商业服务对空间图形和数据库进行两级封装，根据确定的商业规则和商业策略，由空间数据库引擎向数据服务器转发 SQL 指令，并接收返回结果，发往服务前端。

在服务器端，通过空间数据库引擎处理模块，对存入 SQL Server 中的空间图形数据进行访问。服务器在本地执行全部的空间查询和数据提取、存入工作，将满足搜索条件的数据在服务器缓冲存放并经网络发回到商业服务。缓冲处理收集大块数据，然后将整个缓冲区中的数据发往商业服务，而不采用一次只发一条记录的方式。采用这种在服务器端处理并缓冲的方法主要目的是为了提高效率，降低网络传输压力。

在实际开发应用中，空间数据库引擎可以采用两种处理方式(或称为协同处理方式)，处理既可在客户端进行，也可在服务器端进行。对于一些耗时的计算操作，可选在客户端处理。

5. SDG 应用实例

采用空间数据库引擎 Spatial Database Guide(SDG)软件,浙江大学省资源与环境信息系统(GIS)重点实验室开发了一套土地产权产籍管理系统(ReGIS),实现了在 SQL Server 关系型数据库中对空间图形数据进行存取、拓扑运算和拓扑分析操作。图1-6即是该系统的一个界面。系统对地理实体和数据库连接分别进行封装,地籍图形数据通过空间数据库引擎 SDG 存入 SQL Server 数据库中,系统实现了对土地产权产籍管理工作全过程的动态管理。对宗地的变更历史可以进行快速浏览,对于已成为历史状态进入历史库的数据,也可方便地进行图形属性双向互查。系统可以按变更的历史顺序依次撤消,将最后一次认可的历史数据回退到现时库中(参见本章第一节)。此外,系统还提供了图形编辑、输出、数据自动备份、系统修复、互联网信息查询、系统维护及安全保护等功能。在国内一些省、市县的多年应用表明,空间数据库引擎 SDG 运行稳定,性能良好,易于开发,适用面广。在该室承担的水利等其他 GIS 应用领域也得到了很好的应用,替代了昂贵的进口同类产品。综合表现采用该空间数据库引擎的地籍管理系统,如图 7-33 所示。



图7-33 采用自主开发的空间数据库引擎的地籍管理系统

§7.5 组件式 GIS

面向对象程序设计语言和面向对象程序设计方法,已经渗入到计算机软件科学各个领域。随着软件科学的不断发展,新的应用系统越来越复杂,尤其是近几年 Intranet/Internet 的飞速发展,使软件应用置身于更加广阔的环境中,从而对应用软件提出了更高的要求,使软件设计更加困难。在这种形势下,20 世

纪 90 年代诞生了一种新的软件技术，即：组件式（COM）软件技术，组件化程序设计思想得到了迅速的发展，成为当今软件开发技术的主要潮流。

组件式技术的出现为传统 GIS 面临的多种问题提供了全新的解决思路。GIS 软件走过了几十年的发展历程，传统 GIS 虽然在功能上已经比较成熟，但是由于这些系统多是基于多年前的软件技术开发的，受当时技术条件的限制，存在许多固有的缺陷。这些系统基本上属于独立封闭的系统；同时，由于用户需求不断增加，功能一再扩充，GIS 软件变得日益庞大。其结果是用户难以掌握，费用昂贵，阻碍了 GIS 的普及和应用。

组件式 GIS 是 GIS 技术与组件式技术的结合，它使传统 GIS 发生革命性的变化。以往一直是由软件开发商提供全部系统或者具有二次开发功能的软件（内嵌开发语言环境，不能脱离平台），现在除了提供功能强大的集成式 GIS 平台外，还提供丰富的组件工具，由用户根据需要，采用高级开发语言自行开发。可以肯定，组件式 GIS 技术将给整个 GIS 技术体系和应用模式带来巨大影响。

本节首先介绍组件式软件技术。在此基础上，再讲解组件式 GIS 基本概念及特点；最后介绍当今几种主要组件式 GIS 软件平台，并对其功能进行比较分析。

7.5.1 COM 基本概念及实现技术

1. COM 基本概念

COM 是组件对象模型（component object model）的简称。按照组件化程序设计的思想，复杂的应用程序被设计成一些小的功能单一的组件模块，这些组件模块可以运行在同一台机器上，也可以运行在不同的机器上，甚至可以运行在相隔遥远的两台机器上。在理想的情况下，每台机器的运行环境可以不同，甚至可以是不同的操作系统。为了实现这样的应用软件，组件程序和组件程序之间需要一些极为细致的规范，只有组件程序遵守了这些共同的规范，软件系统才能正常运行。为此，国际上 OMG 对象管理组织和 Microsoft 公司分别提出了 CORBA（common object request breaker architecture，公共对象请求中介体系结构）和 COM（component object model，组件对象模型）标准。目前 CORBA 模型主要应用于 UNIX 操作系统平台上，而 COM 则主要应用于 Microsoft Windows 操作系统平台上。Microsoft 提出的组件对象模型标准不仅仅体现了组件化的程序设计思想，在实际上它也采用了面向对象程序设计方法。

简单说来，COM 是一种以组件为发布单元的对象模型，这种模型使各软件组件可以用一种统一的方式进行交互。COM 既提供组件之间进行交互的规范，也提供实现交互的环境，组件对象之间交互的规范不依赖于任何特定的语言，COM 也可以是不同语言协作开发的一种标准。

从起源上讲，COM 标准是在 OLE(object linking and embedding，对象链接和嵌入)技术发展过程中产生的。OLE 技术以 COM 规范为基础，充分发挥 COM 标准的优势，使 Windows 操作系统上的应用程序具有极强的可交互性。如果没有 OLE 的支持，Windows 操作系统则会逊色很多。但是，COM 规范并不局限于 OLE 技术。实际上，OLE 技术只是 COM 的一个应用而已。这些年网络技术飞速发展，OLE 技术在进行网络互连时显示出了很大的局限性，而 COM 则表现出了极强的适应能力。继 OLE 之后，Microsoft 又推出了一系列以 COM 为基础的技术，并统称为 ActiveX 技术。

COM 规范的提出，奠定了 OLE 与 ActiveX 的技术基础。COM 不是一种面向对象的语言，而是一种二进制标准，它定义了组件对象之间基于这些技术标准进行交互的方法。COM 所建立的是一个软件模块与另一个软件模块之间的链接，当这种链接建立之后，模块之间就可以通过称之为对象接口(interface on object)的机制来进行通信，进而实现 COM 对象与同一程序或者其他程序甚至远程计算机上另一个对象之间进行交互，而这些对象可以是使用不同的开发语言，以不同的组织方式开发而成的。COM 定义了一种基础性接口，这种接口为所有以 COM 为基础的技术提供了公共函数。COM 允许组件对其他组件开放其功能调用，既定义了组件如何开放自己以及组件如何跨程序、跨网络实现这种开放，也定义了组件对象的生命周期。

2. OLE 技术概念

OLE 是 object linking and embedding，即“对象链接与嵌入”技术的简称。OLE 不是一个软件，实际上是一组技术。它包括 OLE 包容器、OLE 服务器、就地激活或可视编辑、跟踪、拖放和菜单融合等。这些技术共同创建了全新的 Windows 应用程序。对用户来说，OLE 最明显的使用价值，是在诸应用程序中链接与嵌入文档，它是以文档为中心的计算新概念。然而除了为用户提供新的计算观念以外，OLE 还为开发人员提供了许多新的服务。OLE 是 Microsoft 公司当初为解决 Windows 下应用程序间的通信问题而提出的。但是到了今天，OLE 已演变为一种协议或规范，成为软件构成技术的基础。

最初发布的 OLE 版本只是瞄准复合文档，但在后续版本(OLE2.0)中，导入了 COM。COM 是应 OLE 设计者的需求而诞生的，其基本的出发点是想让某个软件通过一个通用的机构为另一个软件提供服务。因此，COM 的第一个使用者是 OLE2。OLE2 中的新技术包括 Windows 对象的链接、嵌入、就地激活和可视编辑，以及组件对象、结构化存贮、复合文件、统一数据传送、拖放(drag-and-drop)支持和复合文件技术等。在 ActiveX 概念提出以前，OLE 还包括 OLE 控件(现在称 ActiveX 控件)和 OLE 自动化(现在称自动化，在 OLE 和 ActiveX 中都得到支持)。一个 OLE 控件是一个可重复使用的软件组件。OLE 控件可以将

许多其他OLE对象包装在一个包中。这个包可以反映某些特定的功能,如统计图和多媒体等,并可以直接嵌入支持OLE的应用中。OLE对象具有特征和方法,可以通过对它们的编程来控制对象的外观、行为以及相互作用。

对象链接(linking),就是在一个应用程序中的数据和另一个应用程序中的数据之间建立动态链接。例如,可以把电子表格、图表链接到字处理文档中,如果图表数据改变了,则链接到字处理器上的图表也发生相应变化。对象嵌入(embedding)比链接更进了一步。在上例中把电子表格、图表链接到字处理文档中,实际上在字处理器中存贮了图表的一个拷贝。双击这个嵌入的图表,就可以启动了电子表格程序的一个实例(instance),在操纵电子表格中的图表数据的同时就能看到该图表。当退出电子表格应用程序时,字处理文档中的信息随之更新。

3. DCOM基本概念

DCOM是Distributed COM,即分布式COM的简称。早期的COM技术不具备跨计算机的远程调用能力,通过通用接口操纵其他对象的功能,仅局限于同一计算机的不同应用程序之间。例如,Microsoft Visual Basic(VB)可以通过COM通信机制控制和操纵同一计算机中安装的Microsoft Excel的一个拷贝,但不能直接执行其他计算机上的Excel。后来的COM标准增加了保障系统和组件完整的安全机制,扩展到分布式环境。这种基于分布式环境下的COM被称作Distributed COM或DCOM。

DCOM使用一种基于标准的远程过程调用,提供了网络透明及通信自动化,可以使运行于不同机器上的对象之间进行无缝互操作(seamless interaction),而且一个对象无须了解另一个对象的位置。分布式对象技术也可以使全局的网络和信息资源看上去像是本地的,这就使用户更容易也更快地访问重要的业务信息。通过分布式COM和远程自动化,用户可以在整个网络内放置和执行部件,而根本无须知道所处理的信息来自什么地方。因此使用DCOM进行开发,不要求接口的使用者(interface consumer)与接口的提供者(interface provider)必须在同一计算机上。值得一提的是,从纯粹的本地操作移植到分布式操作只需要对现有代码进行少量的修改,有时甚至不需要做任何修改。

一旦COM提供跨网络工作的能力,任何不局限于本地激活模式的接口都具备了分布式的能力。当接口的使用者对某个接口提出请求,该接口可以由另一台计算机上的一个运行着的或者即将运行的对象提供,COM内部的分布式机制提供了建立使用者和提供者之间联接的途径,使用者进行的方法调用(method calls)将出现在提供者一端,并在这里执行,任何返回值将被送回使用者。但一些依赖于本地计算机设备的接口不具有这样的能力,比如绘图接口(drawing interfaces),它需要设备上下文句柄(handles to device contexts)作为参数,而设备

上下文句柄是依赖于本地的接口。Windows NT 从4.0版开始支持DCOM,Windows 95系列也自1996年起提供了支持。在这两种操作系统中,DCOM都包含一组附加的动态链接库(DLL),以及一些同时提供本地和远程COM能力的实用工具。

4. ActiveX 技术

ActiveX 最初是从 Microsoft 的 OLE 技术发展而来的,ActiveX 的整体技术是由 COM 构筑的技术集合。它所涵盖的技术包括三方面:ActiveX 控件(ActiveX Controls)、ActiveX 文档(ActiveX Documents)和 ActiveX 脚本(ActiveX Scripting)。

- ActiveX 文档。ActiveX 文档是对传统嵌入对象的扩展,该文档可以在 Internet 浏览器或者其他 ActiveX 容器程序的整个客户区域中显示,并且可以是多页面的。它可以进行传统的菜单融合,不仅支持就地激活编辑,还支持在服务器程序中开窗编辑。

ActiveX 文档可以使用常用的开发语言编程环境(如 Visual Basic)创建 Web 应用,如代码窗口、调试程序、编译器等。它不仅可以在本地运行,也可以在 Internet Explorer 中运行。在 ActiveX 文档中,可以嵌入各种 ActiveX 部件以进行复杂的运算和操作,实现比带有控件的 HTML 页面更为复杂的客户端处理。支持超链(Hyperlink)对象,用户可以在 HTML 页面和 ActiveX 文档之间透明地定位。特别适合于 Internet 环境下的 GIS 应用系统集成。使用 ActiveX 文档可以很方便地完成复杂的 Web 页面,而且易于实现活动客户与活动服务器之间的 Web 对话以及复杂的数据访问。

- ActiveX 控件。ActiveX 控件是对 OLE 控件概念的扩充。它是基于桌面应用的 COM 组件,一种可编程、可重用的基于 COM 的对象,它遵从一定的与客户端交互的 COM 标准,不仅可以用在一般的 ActiveX 容器程序(如 Visual Basic、Delphi、Visual C ++ 、Borland C ++ 等)中,而且可以用在 Internet 的 Web 页面中,Web 页面中的控件通过脚本实现互相通信。

类似于自动化对象,ActiveX 控件通过属性、方法和事件等接口,与包容器和其他控件进行交互。不同的是,ActiveX 控件通常有自己的显示窗口,增加了 OLE 自定义控件不具备的一些特点。

用户可以根据需要开发各种用途的 ActiveX 控件,如数据库访问、数据监视、数据显示、图形显示、图像处理、三维动画等等。采用 ActiveX 控件方式开发 GIS 基础软件带来了全新的 GIS 概念。在这种方式中,GIS 基础软件以一组分别实现各种功能的 ActiveX 控件提供给用户,用户把这些 GIS 组件嵌入不同的开发环境中实现各种各样的应用系统。

基于组件的应用开发,其方法与组装房子一样,可以用已制作好的组件部件来构筑应用。

ActiveX 控件通过标准化的交互功能，可以在多个不同的上下文中使用同一个控件。ActiveX 控件基本上都是和 Windows 捆绑在一起，以二进制机器代码发放的。随着 ActiveX 控件的推广，ActiveX 控件的开发工具逐日增加。由于 ActiveX 不依赖于语言环境，传统的开发工具基本上都能构筑和配备 ActiveX 控件。最常用的开发工具有 Delphi、PowerBuilder 以及 Visual Basic、Visual C ++、Visual J ++ 等

7.5.2 组件式 GIS 基本概念及特点

1. 组件式 GIS 的基本概念

组件式 GIS 是 GIS 技术与组件式技术的结合。其基本思想是：把 GIS 的各种功能模块进行分类，划分为不同类型的控件，每个控件完成各自相应的功能；各个 GIS 控件之间，以及 GIS 控件与其他非 GIS 控件之间，可以方便地通过可视化的软件开发工具集成起来，形成满足用户特定功能需求的 GIS 应用系统。控件如同一堆各式各样的"汽车部件"，他们分别实现不同的功能（包括 GIS 和非 GIS 功能），根据需要把实现各种功能的"部件"搭建起来，就构成一辆能够完成各种功能的"信息汽车"（应用系统）。

组件式 GIS 系统与传统的各类 GIS 系统相比具有许多特点。把 GIS 的功能适当抽象，以组件形式供开发者使用，将会带来许多传统 GIS 工具无法比拟的优点。在实际工作中，有些应用围绕地图展开，而在其他一些应用中，地图只是其中一部分。此时，应用开发人员迫切需要一种制图与 GIS 功能组件，而不是最终的应用软件来定制或扩展已有的应用。一方面，开发人员希望建立的应用能与通用的桌面软件产品（如文字处理器、数据库等）一起使用；另一方面，希望保留以前的开发成果。对于开发工具，则希望基于一个标准的开发环境，这样可以不必再学习新的编程语言。组件式 GIS 正是为了满足这种需要而开发出来的。

为了进一步体会组件式 GIS 的优势和意义，下面较详细地分析一下传统 GIS 的问题。

2. 传统 GIS 存在的问题

传统的 GIS，主要指的是以 ARC/INFO 和 GenaMap 为代表的功能集成式软件和以 MGE 为代表的模块化软件。这些 GIS 软件平台基本上是基于较早期的软件技术体系设计开发的，在很大程度上已经限制了 GIS 软件进一步的发展和应用。传统 GIS 技术面临严峻挑战，主要表现在以下几个方面：

- 开发负担重　基础软件开发负担过重是传统 GIS 所面临的重要问题之一。为满足各个应用领域在性能和功能上不断提出的更高要求，GIS 基础软件开发者必须不断扩充软件，使其变得庞大臃肿，负担过重。此外，为了使 GIS 软

件成为一个完整系统，开发者除了在软件中实现 GIS 的各种功能与算法外，还不得不做许多与 GIS 底层技术本身无关的工作。其中，用户界面就是较为突出的一例。不管是菜单、工具条、对话框等图形界面的设计，还是宏语言解析的实现，都在 GIS 软件设计和物理实现过程中占有相当比例的工作量。

另外，有些功能，如统计图表制作、数学统计分析、数据库管理等，在其他的专业软件中已经相当完善。但目前的绝大多数商业 GIS 软件中还是不得不“自成体系地”提供这些功能，导致了 GIS 软件与其他软件之间的大量重复劳动，也给使用者增加了学习和掌握 GIS 软件的负担。因此，迫切需要一种全新的软件技术体系，提高程序的可重用性，使 GIS 软件开发者有更多的精力投入到 GIS 本身的技术研究和功能开发中去，避免在开发用户界面和关系数据管理等非 GIS 功能方面花费大量精力。

● 集成困难。传统 GIS 软件封闭、自成体系的结构，使 GIS 很难与应用模型、MIS(管理信息系统)或 OA(办公自动化)等实现高效、有机的集成。随着时间的推移，GIS 软件应用逐步走向综合化，当今 GIS 应用系统可能要集成多媒体、MIS 和 OA 等系统功能，在网络上与无线通讯技术相结合，与 Internet 连接，以及与虚拟现实结合等。传统 GIS 将在这些方面都面临越来越多的困难。

● 二次开发语言复杂。开发语言也是传统 GIS 一直存在的问题之一。大多数 GIS 软件都提供一套自成体系的二次开发语言。例如 Arc/Info 的 AML、MapInfo 的 MapBasic、MicroStation 的 MDL 等等。用户如想在其基础上进行二次开发，不得不学习其语言独特的语法结构、流程控制以及大量的功能函数，才能进行 GIS 应用系统的编码实现。这不仅增加了开发者的负担，也增加了使用者的负担。

● 普及困难。难于普及是阻碍 GIS 应用推广和进一步发展的绊脚石。在当今的社会经济环境中，日常涉及的各种信息的 85% 在一定程度上与地理位置相关。因此，GIS 应当成为一个服务大众的工具。GIS 发展至今，尽管有不少比较成熟的软件系统，但这些软件大都功能繁杂，且系统庞大，非专业技术人员不容易熟悉和掌握这些软件。另外，大多数用户往往只需要 GIS 软件中的一部分功能，但仍不得不为那些他们并不需要的功能花费大量的资金和精力。这些问题给 GIS 技术的普及和大众化带来了困难。

综上所述，在信息技术日益复杂化和综合化的今天，传统 GIS 软件在软件开发、系统集成、用户学习和应用等方面，面临越来越严峻的挑战。因此，迫切需要一种新型的 GIS 软件技术体系，以满足日益提高的 GIS 应用需求，并跟上软件技术发展的潮流。组件式 GIS 技术正是这样一种全新的 GIS 软件技术体系。GIS 技术进入了组件式 GIS 阶段。

3. 组件式GIS的特点

组件式GIS以组件式软件技术作为重要基础，是面向对象技术和组件技术在GIS软件技术开发中的应用。COM GIS控件与其他软件通过标准接口进行通信，实现跨程序、跨计算机、跨网络的分布式操作。

新一代的组件式GIS大多是ActiveX控件或其前身OLE控件。组件式GIS能够使GIS功能嵌入到其他（非GIS）软件中去，或者将其他软件功能引进到GIS软件平台上来，从而使GIS技术与其他软件技术的集成成为现实的可能，这些都体现了组件式GIS的独特优势。组件式GIS为新一代GIS应用提供了全新的开发工具。同传统的GIS相比较，具有多方面的特点，包括无缝集成、跨语言使用、易于推广、开发简捷、使用方便、成本低、可视化界面设计以及Internet应用等。

- 集成灵活、价格便宜。传统GIS结构的封闭性往往使软件本身变得越来越庞大，不同系统的交互性差，系统的开发难度较大。在组件模型下，各组件都集中地实现与自己最紧密相关的系统功能，用户可以根据实际需要选择所需的控件。组件式GIS提供空间数据的采集、存贮、管理、分析和模拟等功能，至于其他非GIS功能（如关系数据库管理、统计图表制作等）则可以使用专业厂商提供的专门组件，有利于降低GIS软件开发成本，最大限度地降低用户的经济负担。组件化的GIS平台集中提供空间数据管理能力，能以灵活的方式与数据库系统连接。在保证功能的前提下，系统表现得小巧灵活，而且价格便宜。这样，用户便能以较高的性能价格比开发出满足需求的GIS应用系统。

- 采用通用开发语言集成。传统GIS往往具有独立的二次开发语言，不仅加重用户和应用开发者的学习负担，而且使用系统开发语言进行二次开发，限制比较多，难以处理复杂问题。而组件式GIS建立在严格的标准之上，GIS应用开发者不必掌握额外的GIS开发语言，只需熟悉基于Windows平台的通用集成开发环境，以及GIS各个控件的属性、方法和事件，就可以实现GIS的基本功能函数，完成应用系统的开发和集成。目前，可供选择的开发环境很多，如Visual C++、Visual Basic、Visual FoxPro、Borland C++、Delphi、C++ Builder、Power Builder等，都可直接成为组件式GIS的优秀开发工具，各自的优点都能够得到充分发挥。这与传统GIS专门性开发环境相比，是一种质的飞跃，不仅有利于减轻GIS软件开发者的负担，而且增强了GIS软件的可扩展性。

- 强大的GIS功能。新的GIS组件都是基于32位系统平台的，采用直接调用形式，无论是管理大数据的能力，还是处理速度均不比传统GIS软件逊色。目前的GIS组件完全能提供拼接、裁剪、叠合和缓冲区等空间处理能力和丰富的空间查询与分析能力。

- 开发简捷、使用方便。由于GIS组件可以直接嵌入应用系统的开发工具

中，对于广大开发人员来说，这就意味着可以自由选用他们熟悉的开发工具。而且，GIS 组件提供的 API 形式非常接近系统工具的模式，开发人员可以像管理数据库表一样熟练地管理地图等空间数据，无须对开发人员进行特殊的培训。在 GIS 应用系统的开发过程中，开发人员的素质与熟练程度是十分重要的因素。这将使大量的应用系统开发人员能够较快地过渡到各类 GIS 专业应用系统的开发工作中，从而大大加速 GIS 的发展。

• 无缝集成。应用组件式 GIS 技术构造应用系统，只需实现 GIS 自身的功能，其他非 GIS 特色的功能（参见表4－1）则可利用其他现有组件。组件之间的联系则由可视化的通用开发语言，如 VB、VC 来实现。通用开发语言建立了软件的框架，软件的"砖头"由组件实现，通过组件之间的消息传递，组件间互相调用，协同工作，从而实现了系统组分之间的高效、无缝集成。

• 可视化界面设计。可以使用 ActiveX 控件的开发语言几乎都支持可视化程序设计。因此，使用 COM GIS 控件集成应用系统，能可视化地设计系统界面，在窗口上布置按钮、列表框和图片框时，可以立即反馈窗口界面的外观，实现所见即所得的界面设计。传统 GIS 软件进行二次开发则需要反复的猜测和实验。

• 更加大众化。组件式技术已经成为计算机软件开发的标准，用户可以像使用其他 ActiveX 控件一样使用 GIS 控件，使非专业的普通用户也能够开发和集成 GIS 应用系统，推动了 GIS 大众化进程。组件式 GIS 的出现，使 GIS 不仅是专家们的专业分析工具，同时也成为普通用户对地理相关数据进行管理的可视化工具。

4. 组件式 GIS 体系结构

组件式 GIS 开发平台通常设计为三级结构。

• 基础组件。基础组件面向空间数据管理，提供基本的交互过程。处于平台最低层，是整个系统的基础，主要面向空间数据管理，提供基本的交互过程，并能以灵活的方式与数据库系统连接。

• 高级通用组件。高级通用组件由基础组件构造而成，面向通用功能，简化用户开发过程，如显示工具组件、选择工具组件、编辑工具组件、属性浏览器组件等。它们之间的协同控制消息都被封装起来。这级组件经过封装后，使二次开发更为简单。如一个编辑查询系统，若用基础平台开发，需要编写大量的代码，而利用高级通用组件，只需几句程序就够了。

• 行业性组件。行业性组件抽象出行业应用的特定算法，固化到组件中，进一步加速开发过程。以水利三维虚拟现实显示为例。对于水利流域规划应用，除了需要地图显示、信息查询等一般的 GIS 功能外，还需要特定的应用功能，如洪水预警模型、流域规划模型、兴利库容计算模型、洪水淹没算法等。这

些水利专业应用功能组件被封装起来后，开发者的工作就可简化为设置各类参数，以及调用接受数据的方法等

各个GIS控件之间，以及GIS控件与其他非GIS控件之间，可以方便地通过可视化的软件开发工具集成起来，形成最终的GIS应用。控件如同一堆各式各样的积木，分别实现不同的功能（包括GIS和非GIS功能），根据需要把实现各种功能的"积木"搭建起来，就构成应用系统。

传统GIS软件与用户或者二次开发者之间的交互，一般通过菜单或工具条按钮、命令以及GIS二次开发语言进行。组件式GIS与用户和客户程序之间则主要通过属性、方法和事件进行交互，如图7-34所示。

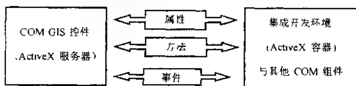


图7-34 COM GIS与客户程序之间的交互

属性（properties）。指描述控件或对象性质（attributes）的数据，如backcolor（地图背景颜色）等。可以通过重新指定这些属性的值来改变控件和对象性质。在控件内部，属性通常对应于变量（variables）。

方法（methods）指对象的动作（actions），如show（显示）、addlayer（增加图层）、open（打开）和close（关闭）等。通过调用这些方法可以让控件执行诸如打开地图文件、显示地图之类的动作。在控件内部，方法通常对应于函数（functions）。

事件（events）指对象的响应（responses）。当对象进行某些动作时（可以是执行动作之前、动作进行过程中或者动作完成后）激发一个事件，以便客户程序介入并响应这个事件。如用鼠标在地图窗口内单击并选择一个地图要素，控件产生选中事件，通知客户程序有地图要素被选中，并传回描述选中对象的个数、所属图层等等有关选择集信息的参数。

属件、方法和事件是控件的通用标准接口，适用于任何可以作为ActiveX包容器的开发语言，具有很强的通用性。目前，可以嵌入组件式GIS控件、集成GIS应用的可视化开发环境很多，根据GIS应用项目的特点和用户对不同编程语言的熟悉程度，可以比较自由地选择合适的开发环境。其中，Microsoft公司的Visual Basic、Visual C++和Borland公司的Delphi功能强大，易于使用，适合大多数GIS应用；而Visual FoxPro等开发环境适合实现数据库管理功能的GIS

应用。表 7－5 为几种可以使用组件式 GIS 控件进行 GIS 应用集成的开发环境比较。

表 7－5 几种组件式 GIS 控件集成的开发环境比较

可视化开发环境	特点及适用范围
Visual Basic、C++ builder Delphi	具有较强的多媒体和数据库管理功能，且易于使用，适合大多数 GIS 应用
Visual C++ Borland C++	功能强大但对编程人员要求很高，适用于编程能力强的用户以及需要编写复杂的、底层的专业分析模型的 GIS 应用
Visual FoxPro Power Builder	数据库管理功能强，适用于建立有大量关系数据的 GIS 应用

7.5.3 几种主要组件式 GIS 平台及其比较分析

当前，地理信息系统软件商纷纷推出或升级已有的组件式 GIS 软件系统，更新的周期越来越快。尤其是几大主流组件式 GIS 平台系统提供商，几乎一年左右就有新的版本推出。其中国际上比较有代表性的系统平台有 ESRI 公司的 MapObjects（简称为 MO），MapInfo 公司的 MapX，Intergraph 公司的 GeoMedia 等。国产组件式 GIS 平台有由武汉吉奥信息工程技术有限公司开发的 GeoMap 和北京超图公司的 SuperMap 平台。

1. MapObjects 平台

MapObjects 是一组供应用开发人员使用的制图与 GIS 功能组件。它由一个 OLE 控件和一系列可编程 OLE 对象组成。开发人员可以利用它在应用程序中添加制图和 GIS 功能；可以在自己熟悉的开发环境中，利用 MapObjects 开发出系统资源占用小的专业应用系统，或在现有的应用中添加 GIS 功能。

MapObjects 是建立在微软的“对象链接和嵌入”（OLE2.0）基础之上的。OLE 是有最广泛支持的面向目标的软件集成技术。用户可以像用砖块盖房子一样利用 MapObjects 的 OLE 组件开发和集成 Windows 应用。MapObjects 提供了超过 45 个可编程 OLE 对象，可以直接插入到许多标准开发环境的工具集中。开发人员可以通过在 VB、VC 等开发环境中建立的属性页来操纵地图，或者通过其他程序化相关对象来控制地图。

MapObjects 的对象分为五大类，即地理匹配对象（address matching objects）、数据访问对象（data access objects）、地图显示对象（map display objects）、几何图形对象（geometric objects）和实用对象（utility objects）。

MapObjects 组件的主要特点包括：

• 兼容广泛的数据格式。MapObjects 支持 ArcInfo COVERAGE, ESRI Shapefiles、ArcSDE 图层和 ImageLayer；栅格图像格式（BMP ArcInfo Grid、ADRG、ASRP、CADRG、CIB、CRP GIF、JFIF、MrSID、NITF、TIFF6、TIFF/LZW、GeoTIFF、ERDAS IMAGINE 和 SVF）等；支持 AutoCAD 13/14/2000 的 DXF、DWG 格式数据；支持矢量产品格式（VPF）数据；支持影像目录（image catalog）管理等。

• 支持 ODBC 和 ADO 访问数据库表格数据源。

• 支持扩展的图形数据及操作。高度数据操作增加了读/写属性，Z 值考虑了所有存贮的几何对象的高度属性，可以利用该属性对特征进行着色；实现了对所有的几何类型数据的量测功能，增加了新的量测属性；允许用户获取和设置任何一点的量测值，即可以得到一个复杂形状（如多边形和线）的每一个顶点量测值；将交叉和交叉属性扩展到了三维环境。

• 具有各种地图投影功能。定义了坐标系统和不同坐标系统之间坐标转换。

• 增强了 GPS 管理功能。MapObjects 支持对点、线和多边形目标对象的追踪。新的方法和属性使得管理和追踪 GPS 动态目标更容易。

• 具有丰富的、用于合并、交叉和缓冲区操作的空间几何库等空间分析功能。

• 提供包含图例、工具条和比例尺的 ArcExplorer 组件原代码。

• 提供对 ArcIMS 支持的新版本的 Weblink 控件和 IMSUtil. dll 库。

• 提供对 ArcSDE 空间数据库引擎的支持。

2. MapX 平台

MapX 是 MapInfo 公司在开发的可视化地图组件 DataMap 基础上，向用户提供的、具有强大地图分析功能的 ActiveX 控件产品。它是一种基于 Windows 操作系统的标准控件，能支持绝大多数标准的可视化开发环境，如 Visual C++、Visual Basic、Delphi、PowerBuilder 等。编程人员在开发过程中可以选用自己熟悉的开发语言，轻松地将地图功能嵌入到应用系统中，并且可以脱离 MapInfo 的软件平台运行。利用 MapX，能够简单快速地在专业应用系统中嵌入地图化功能，增强系统应用的空间分析能力。MapX 采用了基于 MapInfo Professional 的图形内核技术和一致的地图数据格式，根据提供的各种工具、属性和方法，可以实现 MapInfo Professional 所具有的绝大部分地图编辑和空间分析功能。

MapX 提供了 40 多个父类为 Microsoft 基础类（MFC）的 OLE 对象，应用程序员可以使用任何支持 Active X 的程序开发环境开发基于 GIS 的应用程序。从空间数据组织的角度看，MapX 采取点、线、面空间实体和空间索引相结合的方式，通过空间索引，能够以尽可能快的速度查询到给定坐标范围内的空间实体及所对应的数据；MapX 的空间数据采取一种分层存放的结构，用户可以通过

图形分层技术,根据自己的需求或一定的标准对各种空间实体进行分层组合,将一张地图分成不同图层。采用这种分层存放的结构,可以提高图形的搜索速度,便于各种不同数据的灵活调用、更新和管理。

MapX 的主要功能特点包括:

- 地图显示。显示 MapInfo 格式的地图;对地图进行放大、缩小、漫游和选择等操作;或通过 MapInfo 的中间件软件 SpatialWare(一般译为空间数据服务器)或 Oracle 8i 等的空间数据库引擎显示后台地图数据。
- 专题地图。将数据库表中的特定值赋给地图对象的颜色、图案或符号,能用范围值、等级符号、点密度、独立值、直方图和饼图 6 种方式的,基于统计资料的方法来表现专题地图(参见 §6.1.3.2)。
- 图形属性关联。通过简单的点击方式可浏览与地图对象相连的数据信息。
- 数据绑定。地图可通过嵌入 OCX 的容器与数据库相连,并提供了几种不同的数据绑定方式,包括 ZIP Code-level 地理编码法等。
- 地图注解。可提供方向、加亮显示特殊数据,加入文本、符号、表格,使地图信息更加丰富直观。
- 图层控制。显示和控制图层的缩放、使用或创建无缝地图,还可支持一些特殊的应用,如用于实时跟踪活动图层;绘制特殊图形的用户自定义图层等。
- 栅格图像显示。采用栅格图像作为地图的基础图层(背景图层),可使其他图层有一个更细致的背景。
- 自动标记。自动在地图上加入标记,同时标记属性和显示。
- 对象选择。可在地图上拖动鼠标,并在圆、矩形或特定的点上选择一个或多个对象或记录以供分析。
- 对象库。可以使用 Feature Factory 对象,创建、联接或删除点、线、区域图形对象。
- 拖拽工具。使用 MapX 的标准工具或根据需要自己创建的自定义工具,用户可通过点击或拖拽对地图直接进行操作。
- 地图编辑。允许用户添加、修改和删除地图上的点、线、面和文本等特殊对象。
- 投影与坐标系。允许用户调整、选择不同的地图投影和坐标系。
- 远程空间数据服务器。可以访问存贮在 Oracle 8i 和 MapInfo SpatialWare 中的远程地图数据。空间数据服务器,如 SpatialWare、Oracle 8i 等,都提供了先进的查询处理能力,提高了空间数据组织的性能。

3. GeoMedia 平台

GeoMedia 是美国 Intergraph 公司推出的新一代组件式地理信息系统软件,

也是第一个完全采用组件式开发集成的GIS系统。全新的多源数据无缝连接设计思想和先进的数据库管理方式,使得GeoMedia功能强大,有着比传统的GIS鲜明的技术优势。GeoMedia是一个全新的开放式的产品,可以为用户创建纯Windows的应用。所提供的图形、对象和集成能力和增加的许多特别的功能,使得GIS应用系统与其他办公自动化软件能有效地集成在一起,更好的发挥GIS在不同行业和领域中的作用。

GeoMedia的设计思路独特,无论是在总体设计思想,还是在具体实现方法上,都有与其他GIS明显不同之处。GeoMedia独特的设计思想主要有以下几个方面。

• 多源数据无缝集成。GeoMedia目前可以直接读取国际上绝大多数GIS软件的空间数据和属性数据,不需作任何转换,不会丢失信息。它还可以同时把几种不同格式的数据集成在一个坐标系环境中进行空间分析和查询,就像在处理同一种数据一样。这在传统GIS系统中是不能完成的。同时,查询分析的结果与原始数据库动态连接,还可以对分析结果进行再分析。GeoMedia可以直接读取的格式有MGE、FRAMME、ArcInfo、ArcView、MapInfo、SmallWord、SI-CAD、Oracle relational、SQL Server、MS Access、MGE Segment Manager和CAD(包括AutoCAD和MicroStation两种)文件格式。GeoMedia也可输出其他GIS及CAD的格式。

• 先进的数据库管理方式。Intergraph公司在空间数据库的研究和应用中,与Oracle公司合作,开发了内嵌的空间数据库引擎(或称中间件,见§7.4.2),可对Oracle Object、Oracle Relational、SQL Server和Access等直接进行数据读写,不需要中间件,即可将空间数据和属性数据放在数据库的同一记录中,进行统一管理。这个特点为建库、数据更新和历史版本管理,提供了最可靠的高效率的数据管理措施。系统的全部数据都可以由Oracle、SQL Server和Access等商用数据库托管。同时用户可以对存放于大型数据库中的图形、属性数据进行并发操作,并且对大型数据库的历史版本进行管理。

• 功能强大的二次开发环境。GeoMedia使用最新的OLE/COM开发技术,提供了一系列功能较强大的标准开发对象和控件,用户只要利用一种常用开发工具(VB、VC、Delphi或PB等),即可简便地完成应用系统开发工作。基于控件的GeoMedia开发模式简单方便,易于使用,开发技术含量高,能被开发人员在较短的时间内所接受,开发周期短。与其他GIS产品二次开发相比,用户开发出的应用软件功能要强大得多。

• 数据格式标准。GeoMedia没有自己专有的数据格式,该系统的全部数据都由大型商用数据库系统托管,数据标准采用Microsoft Access、Oracle、SQL Server等数据库的标准。因此,基于它的GIS数据库与其他系统可以进行真正

的数据共享和交流、符合开放式GIS(Open GIS)的发展潮流。

• 与Oracle数据库集成。GeoMedia提供了一个GDO服务器,为贮存在Oracle中的空间数据提供服务,并用Oracle原有的优化能力管理空间数据,无须中间件直接访问空间数据库。GeoMedia提供的API,用于从服务器端查询GIS数据,把GIS数据镶嵌到Windows和Windows NT的应用程序中,允许应用程序间数据相互交流和相互操作。

• 全面的高性能数据获取工具。GeoMedia提供了多种方法用于放置和编辑弧段特征,用户可以使用标准编辑命令编辑弧段(如trim、extend、break、split、copy等);具有多边形或线状特征的分离与合并工具,这种工作在数据库内完成,使数据的冗余较小;提供动态验错功能,提高了数据入库质量,避免错误发生;具有智能数据获取功能,可自动确定几何图形特征的顶点、结尾点、原点、中点、栅格图像的拐角点和交叉点,以及线的端点等;具有保证地理特征之间的邻接和拓扑关系的工具,如移动一个街道的交点时,所有与该点相交的特征都移动,保证了地物信息的连接性;编辑多边形特征的边时,与其相邻的边也一起移动,避免产生狭长多边形或缝隙;提供了自动截断线状特征方法;提供了"智能绘图"工具,它包括"版面设计、内容选择、比例关系"等功能。

由于独特的先进的设计思想,GeoMedia具有比较优良的基本功能。GeoMedia与标准关系型数据库、空间数据库协同工作,将空间图形与属性数据存放在同一标准数据库中,提供了数据采集、编辑、分析、报告、输出等高效、实用功能。

在数据采集与编辑方面,GeoMedia添加了完整的GIS数据采集和管理工具,并进行了优化,以提高数据采集和维护的效率。GeoMedia主要的数据采集与维护特点是:灵活的图形特征放置;精确的数据输入;对已知图形特征自动捕捉和断开;对图形特征灵活捕捉;综合的矢量/栅格捕捉;灵活的数据编辑;动态的排对编辑等。GeoMedia具有超强图形编辑能力,它采用Windows图形内核,具有常见的图形编辑功能,支持多文档操作;增加了图形的编辑和出图打印布局(layout)功能,能进行灵活的动态矢量图形捕捉、动态栅格图像捕捉,以及动态排队编辑。能捕捉各种点、线、面间的拓扑状态,使数据采集时能够动态观察拓扑状态、并建立图形拓扑关系,提高工作效率

GeoMedia与办公自动化系统较好地结合,能将项目中的空间数据和属性数据以数据库的形式嵌入到其他办公自动化系统中(如Lotus Microsoft等),使GIS系统和办公自动化系统共享同一数据库。另外,GeoMedia还集成了大多数桌面和办公自动化的应用工具,能直接链接到微软的办公自动化软件,如Word或Excel等系统中,

GeoMedia实现了较复杂的空间分析、查询功能 如对查询集建立缓冲区后,再次对特定区域的数据进行查询;动态分段功能可以查询和分析多种数据

库所创建的分段数据,根据属性数据设计专题地图,从用户的数据中提取有用的信息;能方便的链接多媒体信息,通过添加新的特征来扩充数据集。GeoMedia的分析工具允许用户基于属性数据或空间关系创建复杂查询。这些工具包括9种算子,如完全包含、交迭、接触、空间交集和交迭、空间差别等,形成一套完整的数学算子及其他属性查询的工具。GeoMedia还为用户提供了专门用于空间数据的内部查询算子,它在服务器端运行,充分发挥了服务器的性能。所有的数据是动态连接的,当有变化时都将会动态更新等。

GeoMedia允许将多种空间数据源的数据集成在一个使用同一个坐标系统的用户环境中,以便于访问和查询,操作结果与原始数据可动态链接;通过使用数据服务器访问多种格式的数据源,能够链接世界各地不同物理位置的空间数据库,并能同时对多种数据类型和格式执行分析操作。通过具有的数据综合能力,使用户能将所需数据集成到一个单独的客户环境中。

此外,GeoMedia还具有一定的栅格GIS功能,可集成几乎所有的标准栅格图像格式:CIT、COT、CRL、RGB、RLE、TG4、BMP、GIF、TIF、PCX、CAL、HRF、IGS、JPEG等,并能与各种矢量数据相集成,导出GEOTIFF格式的栅格图像格式,具备精确纠正卫星影像的功能。

4. 几种主要组件式GIS平台功能比较

上面介绍的三种有代表性的组件式GIS平台,即ESRI公司的MapObjects、MapInfo公司的MapX、Intergraph公司的GeoMedia,三者各有特色。下面就它们各自的功能和特点以及不足作一简单的比较,如表7-6所示。表中顺便也列出ESRI公司采用ODE方式的控件(arcedit. ocx、grid. ocx和arcplot. ocx)的特点,但应当指出,ODE控件虽然可以内嵌到通用的可视化编程环境中,但它离不开运行在后台的Arc/Info 7.1.2以上版本环境。从这个意义上说,ODE还不能算作真正意义上的组件式GIS。

表7-6 主要组件式GIS平台功能比较

功　能	MapObjects	ODE	MapX	GeoMedia
显示的地图数据格式	Coverage格式、shp格式数据dwg和DXF格式、多种栅格图像文件	Coverage格式数据和栅格文件	MapInfo的数据格式	MGE、ArcInfo、MapInfo、MS Access、Oracle GDO、和CAD等文件
叠加栅格图像	有	有	有	有

续表

功　能	MapObjects	ODE	MapX	GeoMedia
地图信息查询方式	1　通过鼠标选取特征 2　通过表达式(SQL)查找特征 3　通过空间操作选取特征	1　通过鼠标选取特征 2　通过表达式(SQL)查找特征 3　通过空间操作选取特征	1　通过鼠标选取特征 2　通过表达式(SQL)查找特征 3　通过空间操作选取特征	1　通过鼠标选取特征 2　通过表达式(SQL)查找特征 3　通过空间操作选取特征
图层控制	增加、移走、设置当前层	增加、移走、设置当前层	增加、移走、设置当前层	要素类增加、移走或设置当前层
属性数据绑定	有	有	有	有
对地图的常用操作	放大、缩小、漫游等	放大、缩小、漫游等	放大、缩小、漫游等	放大、缩小、漫游等
专题地图	较弱	没有	有	较强
实时事件跟踪(GPS集成)	有	没有	有	有
用户绘图图层	无	有	有	有
生成和编辑地图对象	较弱	较强	较弱	强
地图标注	有	有	有	有
地图符号化	较弱	较强	较强	强
分析功能	有	有	无	较强
地理编码(地址匹配)	有	无	有	有
可使用的开发语言	VC、VB、PB、Delphi、Access等	VC、Vb等	VC、VB、PB、Delphi、Lotus Notes等	VC、VB、PB、Delphi、等

总起来说，MapObjects 开发简单，性能稳定，适用性也很强；但功能上要明显弱一些。这是因为，ESRI 公司目前的战略思想仍是将强大的 GIS 功能保留在 Arc/Info 和 ArcView 中，而将 MO 定位为一个面向 GIS 查询显示的开发平台。MapX 平台不仅开发非常容易，而且功能也比较强。由于 MapInfo 用户群多，使得很多开发者都乐于使用 MapX 进行应用系统开发。MapObjects 和 MapX 本身并不是完整的 GIS 系统平台，它们主要面向 GIS 开发。但 GeoMedia 不同，它不

仅是组件式GIS开发平台，提供一系列的控件，而且本身也是完整的GIS系统平台，具有数据采集、编辑、管理、制图和空间分析等GIS功能。但GeoMedia的系统稳定性还需进一步提高，开发的总体难度比MapX、MapObjects要大一些。

参考文献

[1] 陈述彭,鲁学军,周成虎. 地理信息系统导论. 北京:科学出版社,2000

[2] 陈述彭. 地学的探索. 北京:科学出版社,1992

[3] 陈述彭. 地球信息科学与区域持续发展. 北京:测绘出版社,1995

[4] 陈述彭. 城市化与城市地理信息系统. 北京:科学出版社,1999

[5] 庄逢甘,陈述彭. 2001·遥感科技论坛. 北京:宇航出版社,2001

[6] 龚建雅. 地理信息系统基础. 北京:科学出版社,2001

[7] 刘南,刘仁义. Web GIS原理及应用——主要Web GIS平台开发实例. 北京:科学出版社,2002

[8] 邬伦,刘瑜,张晶等. 地理信息系统——原理、方法和应用. 北京:科学出版社,2001

[9] 张超,陈丙咸,邬伦. 地理信息系统. 北京:高等教育出版社,1995

[10] 黄杏元,汤勤. 地理信息系统概论. 北京:高等教育出版社,1989

[11] 彭望琭. 遥感数据的计算机处理与地理信息系统. 北京:北京师范大学出版社,1989

[12] 汤国安,赵牡丹. 地理信息系统. 北京:科学出版社,2000

[13] 宋小冬,叶嘉安. 地理信息系统及其在城市规划与管理中的应用. 北京:科学出版社,2000

[14] 陈俊,宫鹏. 实用地理信息系统——成功地理信息系统的建设与管理. 北京:科学出版社,1999

[15] 修文群等. 地理信息系统GIS数字化城市建设指南. 北京:北京希望电子出版社,2001

[16] Michael N DeMers(美),武法东等译. 地理信息系统基本原理. 第二版. 北京:电子工业出版社,2001

[17] 毛锋,孙大路,毕硕本. 模块化地理信息系统环境——MGE基础. 北京:科学出版社,2000

[18] 毛锋. 地理信息系统——MGE方法. 北京:石油工业出版社,1997

[19] 蓝运超,黄正东,谢榕. 城市信息系统. 武汉:武汉大学出版社,1999

[20] 潘正风,杨德麟,黄全义等. 大比例尺数字测图. 北京:测绘出版社,

1996

[21] 陆守一,唐小明等.地理信息系统实用教程.北京:中国林业出版社,2000

[22] 科技部国家遥感中心.地理信息系统与管理决策.北京:北京大学出版社,2000

[23] 李树楷,薛永祺.高效三维遥感集成技术系统.北京:科学出版社,2000

[24] 李志林,朱庆.数字高程模型.武汉:武汉大学出版社,2001

[25] 何建邦,李新通等著.地理信息共享法研究.北京:科学出版社,2000

[26] 毋河海.地图数据库系统.北京:测绘出版社,2000

[27] 三味工作室.MapInfo 6.0 应用开发指南.北京:人民邮电出版社,2001

[28] 潘正风,杨正尧.数字测图原理和方法.武汉:武汉大学出版社,2002

[29] 郝向阳.地图信息识别与提取技术.北京:测绘出版社,2001

[30] 詹庆明,肖映辉.城市遥感技术.武汉:武汉测绘科技大学出版社,1999

[31] 邸凯昌.空间数据发掘与知识发现.武汉:武汉大学出版社,2000

[32] 冯文灏.近景摄影测量——物体外形与运动状态的摄影法测定.武汉:武汉大学出版社,2002

[33] 程建权.城市系统工程.武汉:武汉大学出版社,1999

[34] 陈焕伟编著.土地资源调查.北京:中国农业大学出版社,1998

[35] 林增杰主编.地籍管理.北京:中国人民大学出版社,1990

[36] 李世林,宋英华.大比例尺地形图绘制.北京:测绘出版社,1987

[37] 刘南.地球概论.北京:高等教育出版社,1987

[38] 萨师煊,王珊.数据库系统概论.第三版.北京:高等教育出版社,2001

[39] 苗雪兰,刘瑞新.数据库系统原理及应用教程.北京:机械工业出版社,2001

[40] 刘秋生,何有世.数据库技术.北京:机械工业出版社,2001

[41] 严蔚敏,吴伟民.数据结构.第二版.北京:清华大学出版社,1992

[42] 李春葆.数据结构(C 语言篇)习题与解析.北京:清华大学出版社,2000

[43] 赵东明,刘刚,张岳.数据组织与管理.成都:电子科技大学出版社,1999

[44] 潘爱民.COM 原理与应用.北京:清华大学出版社,1999

[45] 贾永红. 计算机图像处理与分析. 武汉:武汉大学出版社,2001

[46] Jiawei Han, Micheline Kamber(加), 范明, 孟小峰等译. 数据挖掘概念与技术:北京:机械工业出版社. 2001

[47] Adam N A, Gangopadhyay A Database Issues in Geographic Information Systems. Dordreeht: Kluwer Acadamic Publishers, 1997

[48] Andrej V. Interatable and Distributed Processing in GIS. London: Taylor & Francis, 1998

[49] Bailey T C, Gatrell A C. Interactive Spatial Data Analysis. New York: J. Wiley, 1995

[50] Beaumont J R, Gatrell A C. An Introduction to Q-analysis. Norwich: Geo Abstracts, 1982

[51] Buckley A R, Gahegan M, Clarke K. Geographic Visualization. http://www.ucgis.org, 2000

[52] Burrough Peter A. Principles of Geographical Information Systems for Land Resources Assessment (2nd ed.). New York: Oxford University Press, 1989

[53] Burrough Peter A, McDonnell Rachael A. Principles of Geographical Information Systems. New York: Oxford University Press, 1998

[54] Buttenfield B, Gahegan M, Miller H. Geospatial Data Mining and Knowledge Discovery. http://www.ucgis.org, 2000

[55] Clarke K C. Analytical and computer cartography (2nd ed). Englewood Cliffs, N. J.: Prentice Hall, 1995

[56] Craglia M, Couclelis H. Geographic Information Research, Bridging the Atlantic. New York: Taylor & Francis, 1997

[57] DeMers M N. Fundamentals of Geographic Information Systems. New York: John Wiley and Sons, 1997

[58] ERDAS, Inc. Erdas Field Guide (4th ed.). Georgia: ERDAS, Inc., 1999

[59] Fotheringham S, Rngerson P. Spatial Anaysis and GIS. lodon: Taylor and Francis, 1994

[60] Fraser Taylor D R. Geographic Information Systems, Carleton University. Ottawa: Pergamon Press Plc., 1991

[61] Fisher M M, Nijkamp P. Geographic Information Systems, Spatial Modelling and Policy Evaluation. Berlin, Heidelberg: Springer-Verlag, 1993

[62] Fisher P, Unwin D. Virtual Reality in Geography. New York: Taylor & Francis, 2002

[63] Gatrell A C. Distance and Space: A Geographical Perspective New York: Oxford University Press, 1983

[64] Goodchild M F. Geographic Information Science, International Journal of Geographical Information System, 1992, 6(1)

[65] Goodchild M F, Kemp K K. NCGIA Education Activities: The Core Curriculum and Beyond. International Journal of Geographical Information System, 1992, 6(4)

[66] Hearnshaw H M, Unwin D J. Visualization in Geographical Information Systems. New York: Wiley & Sons, 1994

[67] Lloyd R. Spatial Cognition. Dordrecht: Kluwer Acadamie Publishers, 1997

[68] Maguire D J, Goodchild M F. Geographic Information Systems: principle and application. London: Longman, 1991

[69] Mark D M. Geographic Information Science: Critical Issues in an Emerging Cross-Disciplinary Research Domain. http://www.ucgis.org, 1999

[70] Mark D, Egenhofer M, Hirtle S, et al. Ontological Foundations for Geographic Information Science. http://www.ucgis.org, 2000

[71] Muller J C, Lagrunge J P, Weibel R. GIS and Generalization Methodology, and Practice, GIS Data 1. New York: Taylor & Francis, 1995

[72] Moellering H, Clarke D, Cromley R, et al. Analytical Cartography. http://www.ucgis.org, 2000

[73] Molenaar M. An Introdcution to the Theory of Spatial Object Modelling, London: Taylor & Francis, 1998

[74] Samet H. Applications of Spatial Data Structure. Computer Graphics, Image Processing and GIS. Massachusetts: Acldison Wesley Publishing, 1994

[75] Scholl M, Woisard A. Advances in Spatial Databases. Berlin: Springer, 1997

[76] Worboys M F. GIS: A Computing Perspective, New York: Taylor & Francis, 1995

[77] Wright D J, Goodchild M F, Proctor J D. GIS: Tool or Science? Annals of the Association of American Geographers, 1997, 87(2)

[78] University Consortium for GIS. Research Priorities for Geographic Information Science. Cartography and Geographic Information Science, 1996, (23)3

[79] Unwin D. Introductory Spatial Analysis. New York: Methuen, 1981

建议经常浏览的 Web Site

[1] http://www. esri. com

[2] http://www. mapinfo. com

[3] http://www. intergraph. com

[4] http://www. erdas. com

[5] http://www. ucgis org

[6] http://www. opengis. com

[7] http://www. fgdc. gov

[8] http://www. ncgia. ucsb. edu